

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:16:46 ; Search time 20.7407 Seconds
(without alignments)
2109.700 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375

Sequence: 1 MSSLGASFPQIKFDLQFFE.....GDDDDDDGEEDNDMDNSE 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	601	25.3	954	1	S68178 mixed-lineage prot
2	593	25.0	394	2	U000229 mixed-lineage prot
3	581.5	24.5	847	1	AJ53800 mixed-lineage prot
4	565.5	23.8	668	2	JC2363 protein kinase (EC
5	565.5	23.8	888	2	A55318 serine/threonine p
6	560.5	23.6	406	2	T52626 probable mitogen-a
7	560.5	23.6	888	2	UC5399 dual leucine zipper
8	545.5	23.0	328	2	T16747 MAPK delta-1 prot
9	529	22.3	886	2	T48544 MAPK delta-1 prot
10	510.5	21.5	1030	2	P96763 hypotheical prote
11	505	21.3	848	2	B87950 protein F33E2.2 li
12	505	21.3	855	2	T20082 hypotheical prote
13	495	20.8	962	2	T06576 probable protein k
14	488	20.5	821	2	T48400 serine/threonine-p
15	483	20.3	579	2	JC5955 transforming growt
16	480	20.2	829	2	T07406 probable protein k
17	478.5	20.1	1015	2	T00726 probable serine/th
18	475.5	19.9	412	2	T10671 protein kinase hom
19	471.5	19.9	567	2	JC5857 transforming growt
20	471.5	19.9	606	2	JC5856 transforming growt
21	470	19.8	736	2	T05137 protein kinase hom
22	468	19.7	390	2	T01451 protein kinase hom
23	466	19.6	462	2	S29851 protein kinase 6 (
24	460	19.4	963	2	T09911 probable serine/th
25	454.5	19.1	553	2	T04683 hypotheical prote
26	447.5	18.8	407	2	D84635 probable protein k
27	445.5	18.8	546	2	D84555 probable protein k
28	441	18.6	475	2	T12955 probable protein k
29	434	18.3	981	1	F0MVGW gag-abl polypeptel

30	424	17.9	1257	2	T00486 serine/threonine-s
31	421.5	17.7	738	2	F96701 hypotheical prote
32	410	17.3	1130	1	TYHUA protein-tyrosine k
33	407	17.1	842	2	T32258 hypotheical prote
34	405.5	17.1	697	2	A26132 gag-abl-pol polyp
35	405	17.1	1123	2	A39962 kinase-related tra
36	405	17.1	1146	2	B35962 protein-tyrosine k
37	405	17.1	1182	2	A35962 protein-tyrosine k
38	404	17.0	988	2	F86316 protein T1002.13
39	403.5	17.0	1147	2	F86297 hypotheical prote
40	402	16.9	1520	1	TVFPA protein-tyrosine k
41	401.5	16.9	1584	2	T18276 protein-tyrosine k
42	400.5	16.9	1094	2	S49313 protein kinase - s
43	400	16.8	1248	2	B96827 hypotheical prote
44	398.5	16.8	1171	2	T12956 hypotheical prote
45	396.5	16.7	545	2	T05675 hypotheical prote

ALIGNMENTS

RESULT 1

S68178 mixed-lineage protein kinase 2 (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)

C:Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text_change 11-Jun-1999

C:Accession: S68178; J38044; S32468

R:Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps

Eur. J. Biochem. 234, 492-500, 1995

A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of human

A:Reference number: S68178; MUID:96128179; PMID:5536994

A:Accession: S68178

A:Molecule type: mRNA

A:Residues: 1-954 <DOR>

A:Cross-references: EMBL:X90846; NID:9971419; PIDN:CAA62351.1; PID:9971420

R:Kato, M.; Hirai, M.; Sugimura, T.; Terada, M.

Oncogene 10, 1447-1451, 1995

A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase

A:Reference number: J38044; MUID:9549256; PMID:7731697

A:Accession: J38044

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806, 'R', 808-817, 'A', 819-954 <RSS>

A:Cross-references: EMBL:Z48615; NID:9758592; PIDN:CA88531.1; PID:9758593

R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.

Eur. J. Biochem. 213, 701-710, 1993

A:Title: Identification of a new family of human epithelial protein kinases containing ty

A:Reference number: S32468

A:Accession: S32468

A:Molecule type: mRNA

A:Residues: 244-464, 'AQAAGROPHPALWL' <DQ2>

C:Genetics:

A:Gene: GDB:MLK2; GDB:MST

A:Cross-references: GDB:362654; GDB:624810; OMIM:600137

A:Map position: 19q13.1-19q13.2

C:Superfamily: mixed-lineage protein kinase 2, protein kinase homology; SH3 homology

C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein ki

F:23-76/Domain: SH3 homology <SH3>

F:36-364/Domain: protein kinase homology <KIN>

F:104-112/Region: protein kinase ATP-binding motif

F:384-405/Region: leucine zipper motif

F:419-440/Region: leucine zipper motif

F:449-463/Region: basic

F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 25.3%; Score 601; DB 1; Length 954;

Best Local Similarity 35.4%; Pred. No. 4.7e-25;

Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps 10;

QY 10 QIKFDLQFFENCGGSPGVYRAKWIISQDKVAVKKL-LKIER-----EAEIL 57
DB 92 EIPFHELOQEITIGVGFKKYVRLW--RGEVAVKARLDEKDPAYVTAQVCEARLF 149

```

QY 58 SVLSHRIITQFVGVILLEPPNYGIVTEYASLSGLYDINSRSEMDMDHIM-TWATDVAK 116
DB 150 GALQHNIIALRGACLNPPHLCIMVEYARGALSRVLAGRVP-----PHVLVMAVOVAR 205
QY 117 GMYHLMHEAPVKYIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVG 168
DB 206 GMYHLMHEAPVKYIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVG 265
QY 169 TFPWMAPEVYQSLPVSETCTDYGVVLMEMLTREVPFKLEGLQVAMLVENKRELTTP 228
DB 266 TYAMMAPEVYRLSLFSSKSSVMSFGVLLMELLTGEVYREIDALAAYAVANMKLTLP 325
QY 229 SSCPSRPAELLHOCWEADAKKPSFKQIISLESMSNDTSLPDKMSFLHNKAEMCEIE 288
DB 326 STPEEPFARLIEECWDDPHGRDPFGSILKRLVIEQSALFQMPLESFHSLOEDMKLEB 385
QY 289 ATLERLKLERDLSFKEOEL-----KERERRLKWEQKLTQOSNTPLPLPLAAMSEB 341
DB 386 HMFDDLRTEKELKRSHEBELLRAAQEQRFQEQELRRREGRLAREMDIVERELHLMCQL 445
QY 342 SYRESKTESNSAEMSCQITATNGEGHGMNPSLOAMTLMFGDDITSMNKAGAVHSGMO 401
DB 446 SQEKPRVRRKKGNFKSRRLKLRREGSGHSLPS-----GFEH-----K 483
QY 402 INMQAKONSKSTSKRG 419
DB 484 ITVOA-----SPTDKRG 497

```

RESULT 2

```

J00229
mixed-lineage protein kinase 1 - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C/Accession: S32467; J00229
R/Dorc, D.S.; Devereux, L.; Dietzsch, E.; de Kreseer, T.
Eur. J. Biochem. 213, 701-710, 1993
A/Title: Identification of a new family of human epithelial protein kinases containing
A/Reference number: S32467; MUID:93238756; PMID:8477742
A/Accession: S32467
A/Molecule type: mRNA
A/Residues: 1-394 <DO2>
C/Genetics:
A/Gene: GDB:MLK1
A/Cross-references: GDB:141921; OMIM:600136
A/Map position: 14q24.3-14q31
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F/1-269/Domain: protein kinase homology <KIN>
F/1-269/Domain: catalytic <CAT>
F/9-17/Region: protein kinase ATP-binding motif
F/289-310/Region: leucine zipper motif
F/324-345/Region: leucine zipper motif
F/354-368/Region: basic

```

```

Query Match 25.0%; Score 593; DB 2; Length 394;
Best Local Similarity 39.3%; Pred. No. 4.7e-25;
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

QY 15 DLQPFENCGGSGFSGYVRAKMIISODKEVAVK-----KLKIKKEAELISVLSH 62
DB 2 ELTLEELIIGIGGFGKYRAFWIGD--EVAVKARHPDEDEISOTIENVREALKLPAMLK 59
QY 63 RNIIQFVGVILLEPPNYGIVTEYASLSGLYDINSRSEMDMDHIMTWATDVAKGMHYH 122
DB 60 PNIIALRGVCLKEPNCLVMEFARGGLNVLVSGKR---IPDILVMAVOIARGMNYH 116
QY 123 MEAPVKYIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVGTTPMA 174
DB 117 DEAIIVPIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVGTTPMA 176
QY 175 PEVIQSLPVSETCTDYGVVLMEMLTREVPFKLEGLQVAMLVENKRELTTPSSCPRS 224
DB 175 PEVIQSLPVSETCTDYGVVLMEMLTREVPFKLEGLQVAMLVENKRELTTPSSCPRS 224

```

```

DB 177 PEVIASMSKSGSDVMSYGVLLMELLTGEVPRGIDGLRVAYGVANMKLLPISPCBP 236
QY 235 FAEHLHOCWEADAKKPSFKQIISLESMSNDTSLPDKNSFLHNKAEMCEIEATLE 294
DB 237 FAKLMEDCNPPHSPHSPFTNILDQTLTIEEGFFEMPDPKSFHCLQDNKHEIQEMFDL 296
QY 295 KKLERDLSFKEOEL-----KERERRLKWEQKLTQEQ 325
DB 297 RAKEKELRTWEELTRALQKNQEBELRRRQELAE 334

```

RESULT 3

```

A53800
mixed-lineage protein kinase (EC 2.7.1.-) 3 - human
N/Alternate names: protein kinase PTK1; protein kinase SPRK
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A53800; I58395
R/Gallo, K.A.; Mark, M.R.; Seadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
A/Reference number: A53800; MUID:94253068; PMID:8195146
A/Accession: A53800
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-847 <GAL>
A/Cross-references: GB:U07747; NID:9464027; PIDN:AA19647.1; PID:9464028
R/Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Laessle, N.J.
Oncogene 9, 1745-1750, 1994
A/Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain
A/Reference number: I58395; MUID:94239754; PMID:8183572
A/Accession: I58395
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-847 <RES>
A/Cross-references: GB:L13976; NID:9488295; PIDN:AAA59859.1; PID:9488296
C/Genetics:
A/Gene: GDB:MLK3; PTK1; SPRK
A/Cross-references: GDB:134755; OMIM:600050
A/Map position: 11q13.1-11q13.3
C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F/48-100/Domain: SH3 homology <SH3>
F/115-383/Domain: protein kinase homology <KIN>
F/123-131/Region: protein kinase ATP-binding motif
F/403-424/Region: leucine zipper motif
F/438-459/Region: leucine zipper motif
F/468-482/Region: basic

```

```

Query Match 24.5%; Score 581.5; DB 1; Length 847;
Best Local Similarity 37.4%; Pred. No. 4.6e-24;
Matches 132; Conservative 60; Mismatches 122; Indels 39; Gaps 7;

QY 13 FDLQPFENCGGSGFSGYVRAKMIISODKEVAVK-----DKVAVKLLKIKKEAELISVLS 61
DB 114 FOELRLIEVYIGIGGFGKYRSGMRGEVAVKARQDPDEDIV-TAESVROEARLPAMLA 172
QY 62 HNNIIQFVGVILLEPPNYGIVTEYASLSGLYDINSRSEMDMDHIM-TWATDVAKGMHY 120
DB 173 HNNIIALKAVCLKEPNCLVMEYAAAGPLSRALAGRVP-----PHVLVMAVOIARGMNY 228
QY 121 LMEAPVKYIHRDLKSRNVV-----IAADGV-----LKICDFGASRFNNHTTHMSLVGTTFW 172
DB 229 LICEALVPIHRDLKSRNVV-----IAADGV-----LKICDFGASRFNNHTTHMSLVGTTFW 288
QY 173 MEAPVKYIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVGTTPMA 232
DB 289 MEAPVKYIHRDLKSRNVV-----AADGVLTICDFGASRFNNHTTHMSLVGTTPMA 348
QY 233 RSEFAELLHOCWEADAKKPSFKQIISLESMSNDTSLPDKNSFLHNKAEMCEIEATLE 292
DB 349 BPEAQMDACMQDDPHRRDFAISIIQQLEALBAQVLRMPBRDSFMSMGCKREIOGLPD 408

```


Db 643 EIMMEDDQIGERIGIGSGEYVRAEW--NGTEVAVAKKRLDDFGSDALTQKSEIBIMLR 700

Qy 60 LSHRNIIIOFYVILPEPPYGVITEVYASIGSLYDIYNSNRSEEMDMHMTATYAKGM 119

Db 701 LRHPNVVLFMGAVVTRPPNFSILTEFLPRGSLYRLRH--RPNHQDEKRRMRVALYAKGM 759

Qy 120 YLHNEAPKVIHRRDLSKNVVIADGVGLKICDPGASRPHNHT--THMSLVTFPPMAPEV 177

Db 760 YLHSHPL-VHRRDLSKNVLDKKNVAVKCDPGLSRKHHTHYSSSTAGTPENMADEV 818

Qy 178 IQSLPVSETCDTYSYGVVLMELTREVPFKGLEQLQVLMVLEKNERLTIPSSCPSPAE 237

Db 819 LRNEPANKCOVYSGVILWELATSRVPMKGNPMQVGVANGFQNRRLIEDDIDLTVAQ 878

Qy 238 LLHQCEW 244

Db 879 IIRECWQ 885

RESULT 10

F96763

hypothetical protein F25P22.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: F96763

R:Rheologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salbeck, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A6141; MUID:21016719; PMID:11130712

A:Accession: F96763

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1030 <STO>

A:Cross-references: GB:AB005173; NID:g6692730; PID:AAF24836.1; GSPDB:GN00141

C:Genetics:

A:Gene: F25P22.8

A:Map position: 1

Query Match 21.5%; Score 510.5; DB 2; Length 1030;

Best Local Similarity 38.0%; Pred. No.3.7e-20;

Matches 109; Conservative 54; Mismatches 99; Indels 31; Gaps 7;

Qy 10 QIKEDDIOFPENCGGSGFGSYVRAKMTISQDEKAVAKLL-----KIEKEAILSV 59

Db 742 EILWEIITVGRIGIGSGEYVRAKMTISQDEKAVAKLLDDLTGEALEPFRSEVRIMKK 799

Qy 60 LSHRNIIIOFYVILPEPPYGVITEVYASIGSLYDIYNSNRSEEMDMHMTATYAKGM 119

Db 800 LRHPNVVLFMGAVVTRPPNFSILTEFLPRGSLYRLRH--RPNHQDEKRRMRVALYAKGM 858

Qy 120 YLHNEAPKVIHRRDLSKNVVIADGVGLKICDPGASRPHNHTTH--SLVGTPEPMAPE 176

Db 859 YLHSCNPM-IYHRDLSKNVLDKKNVAVKCDPGLSRN-KISTYSSSTAGTAEWMAPE 916

Qy 177 VIQSLPVSETCDTYSYGVVLMELTREVPFKGLEQLQVAMVLEKNERLTIPSSCPSPFA 236

Db 917 VLIRNEPAPKCDVYSGVILWELFTLQOPWGMGNPMQVGVANGFQHRRLTIDPFDVPAIA 976

Qy 237 ELLHQCEWADAKKRPSEKQIISILEMS-----NDTSLP 270

Db 977 DLISKCWQTDLSKRPSEFAEIMASLRLQKPVGTGSIIPRPVSSSLP 1023

RESULT 11

887950

protein F3J3E2.2 [imported] - Caenorhabditis elegans

[illegible]

```

A:Accession: J121703
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-855 <W12>
A:Cross-references: EMBL:284574; PIDN: CAB06544.2; GSPDB: GN00019; CESP: F33E2.2
A:Experimental source: clone F33E2
C:Genetics:
A:Gene: CESP: F33E2.2
A:Map position: 1
A:Introns: 47/2; 213/2; 263/2; 366/3; 494/1; 547/3; 608/1; 826/3

Query Match      21.3%; Score 505; DB 2; Length 855;
Best Local Similarity 30.4%; Pred. No. 6e-20;
Matches 145; Conservative 69; Mismatches 159; Indels 104; Gaps 14

QY 10 QIKEDDLOFFENCGGSGFSGYVRAKMIISODKEAVAKLLIEKEAEI--LSVLSHRNIQ 67
DB 56 EIPFDALSELMLGSGGAVFRQG--LENRVAVKYNQL-KETELIKHLRLHRNQIIE 112
QY 68 FYCVILPEPNYGVITEYASLSGLYIYINSNSEMDMHIMTADVAKGHYLLMEAPV 127
DB 113 FLGVCSKSPCYCIWVEYCSKQQLCTVLKSRNITREL--FAQWVEIADGMHYLHQN-- 167
QY 128 KVIRHDLRSRVVITADVVKICDGFARFHN--HTHMSLVGFPPMAPEVIOSLPSE 185
DB 168 KVIHRDLSPNLLISAESIKICDGTSHMQKMDSTNMFSGVISMVAPETIKQPCNE 227
QY 186 TCDTYSYGVVLWMLTREVPFKGLEGLQVAMLVENKSRILTIPSSCPSSFALLHQCEA 245
DB 228 KVDVYSFQVVLWEMLTRETPYANIQAQMIIFGVGINILSLPMBEAPRGVILLIQCLSQ 287
QY 246 DAKRPSKQI-----ISLMSNDISLPD-- 271
DB 288 KGRNRPFSHTRHWEIKFPELTFEETSEWQLAMDYREFAKCIQYPSVTVRDHGGPKSA 347
QY 272 -----KNSFLHNKAEMRCEIATELTERLKKLERLSPFEGCLKREKRLKWEQKL 322
DB 348 FAMEBEIQRKHEQLNHLKDIRNMYEMKIKRTNKMVDKIQGCTELKKESELAWEKDL 407
QY 323 TEQ-----SNTP--LLPLAARMS--EESYFESKTEES-----NSAEMSQIATSNGEQH 369
DB 408 TEREGMHNQNSPKVAAPRAQIRGVPNNGYDDMSDEVDQPCRGSPYCSNTHSS----- 462
QY 370 GMAPSLQAMMLMGFDIFSNKAGAVMHSQMOINQAONSSKTTSS--KRRGKRYN 423
DB 463 -----SGVSSPFSRSSSSRSASAGOOTRSEGAN 491

RESULT 13
T06576
probable protein kinase TC2R2 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C:Accession: T06576
R:Hackett, R.M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z15770
A:Accession: T06576
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-892 <HAC>
A:Cross-references: EMBL:AJ005077; NID:e1296722; PIDN: CAA06334.1; PID: e1296723
A:Experimental source: cultivar Alisa crataeg
C:Genetics:
A:Gene: TC2R2

Query Match      20.8%; Score 495; DB 2; Length 982;
Best Local Similarity 38.9%; Pred. No. 2.4e-19;
Matches 110; Conservative 49; Mismatches 104; Indels 20; Gaps 7

QY 10 QIKEDDLOFFENCGGSGFSGYVRAKMIISODKEAVAKLL-----KIEKEAIIISV 59
DB 695 EIPWEDLVIGRIGIGSYGEVYHADW--NGTEVAVKKEFLDQDFSGAALAEFRKRVIMR 752

```

QY 60 LSHENIIQFYGVILPEPPYGVITEYASLGSLVDYINSNSEEMDMHITWATDVAKGM 119
D 753 LRHNHYVAFKALITRRPPLSLITTEFLPRGSLYRIITRRHF-QIDERQKIMALDVAKGM 811
QY 120 YLHMEAPKVIHRDLKSRNVVIAADGVLKICDFGASRFPHNT--THMSLVGFPMNAPEV 177
D 812 CLHFSNPT-IYHRDLKSNPLLVDTDMNVKVCDFGLSRLKHNHFLSKSTAGTEEMNAPEV 870
QY 178 IQSLPVEPTCCTTYGVGVLMELTREVPEKLEGLOVAMLVYEKNRRLTIPSSCPRSFAE 237
D 871 LRNPSPNEKCIYSGFVLMEATLRLPMSGNNPQGVAGVGPQNRLEIPKELDPVAR 930
QY 238 LTHOCWEADAKKRPSFKQIISILBSMSNDTSLP---DKCNSFL 277
D 931 IIMECQWTDPLRPSFQLTVALTRPLQR-LVIPAVDQINSRL 972
RESULT 14
T48400
serine/threonine-protein kinase ctrl - Arabidopsis thaliana
N/Alternate names: protein P17C15.150
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #next_change 17-Nov-2000
C/Accession: T48400; A45178
R/Bevan, M.; Pohl, T.; Weizengeger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A/Reference number: Z24492
A/Accession: T48400
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-821 <BEV>
A/Cross-references: EMBL:AL162506
A/Experimental source: cultivar Columbia; BAC clone P17C15
R/Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.
Cell 72: 427-441, 1993
A/Title: CTRL, a negative regulator of the ethylene response pathway in Arabidopsis, encodes
A/Reference number: A45178; MUID:93161417; PMID:8431946
A/Contents: Columbia
A/Accession: A45178
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-468, 470-821 <KIE>
A/Note: sequence extracted from NCBI backbone (NCBIP:124878)
C/Genetics:
A/Map position: 5
A/Introms: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/3; 7
A/Note: P17C15.150
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP
F/549-812/Domain: protein kinase homology <KIN>
F/557-565/Region: protein kinase ATP-binding motif
Query Match 20.5%; Score 488; DB 2; Length 821;
Best Local Similarity 40.7%; Pred. No. 4,7e-19;
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;
QY 9 VOIKRPDILQIFENCGGSGFSGYRAKWIISQKQVAVKYLKIE-----KEAEILS 58
D 544 MDIPWCDLINIEKIGAGSGFYHRAEMHGS--VAVKIIMEDDFAHRAVNEFLREVAIMK 601
QY 59 VLSHRNIIQFYGVILPEPPYGVITEYASLGSLVDYIN--SNRSEEMDMHITWATDVAKG 117
D 602 RLRLPNIVLFGAVATQPNPLSLITTEFLPRGSLYRIITRRHF-QIDERQKIMALDVAKG 661
QY 118 MHYLHMEAPKVIHRDLKSRNVVIAADGVLKICDFGASRFPHNT--THMSLVGFPMNAPEV 175
D 662 MNYLHNRNP-PIVHRDLKSNPLLVDTDMNVKVCDFGLSRLKHNHFLSKSTAGTEEMNAPEV 720
QY 176 EVIQSLPVEPTCCTTYGVGVLMELTREVPEKLEGLOVAMLVYEKNRRLTIPSSCPRSF 235
D 721 EVLDEBENESKDIYSGFVLMEATLRLPMSGNNPQGVAGVGPQNRLEIPKELDPVAR 780

```
QY      236 AELLHQWEADAKRPSFKQIISILESM 263
          |::|:|||||:|
Db      781 AAIEGCTNEPWRKPSFATIMDLRLPL 808
```

RESULT 15

t:transforming growth factor-beta activated kinase (EC 2.7...-.) 1a - human
 C:Species: Homo sapiens (man)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: J05955
 R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
 Biochem. Biophys. Res. Commun. 243, 545-549, 1998
 A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
 A:Reference number: J05955; MUID:98155801; PMID:9480845
 A:Accession: J05955
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-579 <SAK>
 A:Cross-references: DDBJ:AB009356; NID:G2924623; PIDN:BAA5025.1; PID:G2924624
 C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homolo
 C:Keywords: phosphotransferase

Query Match	20.3%	Score 483;	DB 2;	Length 579;
Best Local Similarity	30.2%	Pred. No. 5.9e-19;		
Matches 131;	Conservative 82;	Mismatches 151;	Indels 70;	Gaps 18

```

QY 7 SFOIKEDDLOFENCNGGSGFSGSVYRAKMIQODKEVAKUKLTIKKEAE-----IL 57
Db 27 NFESEIDYKEIIEVEVEVGRGAFVGYCKAKM--RAKDAI--QLESSEBAFIIVEIROL 81
QY 58 SVLSHRNIIQFYGVILLEPPNYGIYTERVASYLSGLYDIYNSNSEE--DMDHITMATDV 114
Db 82 SRVNHPIVATLYGACINP--VCLINVEYAGSLVYVNLHG--AEPLPYTAAHAMSWCLOC 137
QY 115 AKGMHVLHMEAPKVKVHVRDLKSRNVIAADG--VLKICDPCG--ASRPHNHTTMSLVGFPPW 172
Db 138 SQGVAAVLHSMQFALVHRLKRPNNLLVAGGVTLKICDGFACDIOQTMITNK--GSAAW 195
QY 173 MAEEVIOQLPVSETCDTYGVAVLEWMLFREVPFGGLEG--LOYAMVLVERNERTLTPSS 230
Db 196 MAEEVEGSGYSKCOVFSMGIIIMVITRRKRPDEIGSAPRIMW--AVNHGTPPELLKN 255
QY 231 CPSPSEMLHOCHEADAKTRPSEKQIISL-----ESMSNDT- 267
Db 255 LPPIRESLMTRCXSKDPSORPSMEELVYKIMTILMEFPGADEPLQPCQYDEGOSNAT 314
QY 268 ---SLDEKCNSTFLHNAEMRCE--IEHTLERLKKLERDLSFDEOLKEXERLUXMEOKLT 322
Db 315 STGSPMDIASTNTSNKSDTNTMEQVPATNDTIRLSKLLKNOAKQOSESGLSLGAARGS 374
QY 324 EOSNTPLLPLFARMSEESYFESKITEESASAMSCQIATSGEHCINPISLOAMMIMGF 383
Db 375 SVESLSP---PTS-----EGKMSAMMSIEIARIAATT--GNGQPRRRRIODLVTYGT 421
QY 384 --GDIFSANKAGAV 395
Db 422 EPGQVSSRSSSPSV 435

```

Search completed: December 5, 2003, 09:20:58
Job time : 22.7407 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 08:32:56 ; Search time 12.963 Seconds

(without alignments)
1650.637 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375

Sequence: 1 MSSLGASFYQIKFDLQFFE.....GDDDDDDGGEEDNDMDNSE 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	601	25.3	954	1	M3KA_HUMAN	Q02779 homo sapien
2	593	25.0	934	1	M3K9_HUMAN	P80192 homo sapien
3	565.5	23.8	859	1	M3KC_HUMAN	Q12852 homo sapien
4	563.5	23.7	888	1	M3KC_MOUSE	Q60700 mus musculu
5	488	20.5	821	1	M3K7_MOUSE	Q63796 ratcus norv
6	471.5	19.9	579	1	M3K7_MOUSE	Q05609 arabidopsi
7	471.5	19.9	579	1	M3K7_MOUSE	Q05609 mus musculu
8	471.5	19.9	579	1	M3K7_MOUSE	Q05609 mus musculu
9	434	18.3	746	1	ABL_MOUSE	P00521 abelson mur
10	434	18.3	746	1	ABL_MOUSE	P00521 abelson mur
11	405	17.1	1123	1	ABL_MOUSE	P00520 mus musculu
12	405	17.1	1123	1	ABL_MOUSE	P00520 mus musculu
13	402	16.9	1520	1	ABL_MOUSE	P00522 dirosophila
14	401.5	16.9	1520	1	ABL_MOUSE	P00522 dirosophila
15	385.5	16.2	439	1	ABL_MOUSE	P10447 feline sarc
16	385.5	16.2	439	1	ABL_MOUSE	P10447 feline sarc
17	379.5	16.0	536	1	FRK_MOUSE	P27446 xiphophorus
18	378.5	15.9	393	1	FRK_MOUSE	P42685 homo sapien
19	378.5	15.9	393	1	FRK_MOUSE	P42685 homo sapien
20	376.5	15.9	536	1	FRK_MOUSE	P83104 dirosophila
21	374	15.7	410	1	FRK_MOUSE	P06241 mus musculu
22	369.5	15.6	1224	1	ABL_MOUSE	P18161 dirosophila
23	365.5	15.4	535	1	FRK_MOUSE	P03949 caenorhabdi
24	362	15.2	1317	1	PR2_MOUSE	Q02977 gallus gall
25	359.5	15.1	533	1	FRK_MOUSE	Q91717 dirosophila
26	358.5	15.1	587	1	FRK_MOUSE	P39688 mus musculu
27	358.5	15.1	675	1	FRK_MOUSE	P15054 avian sarco
28	358	15.1	822	1	FRK_MOUSE	P51813 mus musculu
29	358	15.1	822	1	FRK_MOUSE	P51813 mus musculu
30	357	15.0	819	1	FRK_MOUSE	P16092 mus musculu
31	357	15.0	819	1	FRK_MOUSE	P16092 mus musculu
32	355.5	15.0	517	1	FRK_MOUSE	P21804 gallus gall
33	355.5	15.0	531	1	FRK_MOUSE	P14234 mus musculu

34	355.5	15.0	531	1	SRC2_XENLA	P13116 xenopus lae
35	354	14.9	497	1	SPK1_DUGTI	P42687 dugesia tig
36	353.5	14.9	532	1	SRC2_CHICK	P00523 gallus gall
37	353	14.9	822	1	FRK_MOUSE	P00523 gallus gall
38	352	14.8	806	1	FRK_MOUSE	P00523 gallus gall
39	351.5	14.8	557	1	SRC2_CHICK	P18460 gallus gall
40	351.5	14.8	557	1	SRC2_CHICK	P18460 gallus gall
41	350.5	14.8	527	1	TXK_MOUSE	P53156 hydra atten
42	350.5	14.8	533	1	FRK_MOUSE	P42681 mus musculu
43	350	14.7	806	1	FRK_MOUSE	P05876 gallus gall
44	350	14.7	1426	1	FRK_MOUSE	P22607 mus musculu
45	349.5	14.7	535	1	SRC_MOUSE	P04412 dirosophila
						P12931 mus musculu

ALIGNMENTS

RESULT 1
M3KA_HUMAN STANDARD; PRT; 954 AA.
AC 002779; Q12761; Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)
DE (Mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MLK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA Sutherland G.R., Simpson R.J.,
RT "Complete nucleotide sequence, expression, and chromosomal
RT localisation of human mixed-lineage kinase 2.";
RL Eur. J. Biochem. 234:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95249256; PubMed=7731697;
RA Katon M., Hirai M., Sugimura T., Terada M.,
RT "Cloning and characterization of MST, a novel (putative)
RT serine/threonine kinase with SH3 domain.";
RL Oncogene 10:1447-1451(1995).
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.,
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
RN [4]
RP CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X90846; CAA62351.1; -
CC EMBL: X48615; CAA88531.1; -
CC PIR: S68178; S68178.

```

DR HSSP: P11362; 1FGK.
DR Genew; HGNC:6849; MAP3K10.
DR MIM; 600137; -.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR GO; GO:0007254; P:JNK cascade; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinaase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50002; SH3; 1.
DR Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; SH3 domain.
FT DOMAIN 2 5 POLY-GLU.
FT DOMAIN 16 81 SH3.
FT NP_BIND 98 360 PROTEIN KINASE.
FT BINDING 104 112 ATP (BY SIMILARITY).
FT ACT_SITE 125 125 ATP (BY SIMILARITY).
FT DOMAIN 222 222 BY SIMILARITY.
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).
FT CONFLICT 465 480 LKLRGEGSHSLPSGF -> AQAGRRQHPQALML (IN REF. 3).
FT CONFLICT 471 471 G -> S (IN REF. 2).
FT CONFLICT 807 807 G -> R (IN REF. 2).
FT CONFLICT 818 818 V -> A (IN REF. 2).
SQ SEQUENCE 954 AA; 103623 MW; 538F4AA559B0ABA CRC64;

Query Match 25.3%; Score 601; DB 1; Length 954;
Best Local Similarity 35.4%; Pred. No. 1.6e-31;
Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps 10;

QY 10 QIYFDLQFFENCGGSGFSGVYRAKWSQDKYAVAKKL-LKIKK-----EAEIL 57
DB 92 EIPFHELOEELIIGVGFGKVRALM--RGEVAVAAARLDPEKDPAYVAQVCEARLF 149
QY 58 SVLSHNIIOFYVILPEPVYGIYTERYASLSGLYDYNRSRSEMDMDHIM-TWADVAK 116
DB 150 GALQHNIITLAKACINPRLCLVMEYARAGALSRLVLAGRRV---PHYLVMAVQVAR 205
QY 117 GMHYLHMEAPVKYIHDLKSRNVI-----ADGVLTICDGFASRFPHNTTHMSLVG 168
DB 206 GMNVLHNDAPVPIIHDLSKINILILEAJENHNLADTLKTFDGLAREWHKTTKSAAG 265
QY 169 TTPWMAPEVYIQSLPVSECTDYGVVWMLTREVPRFKQBELQVAMLVENKELTTP 228
DB 266 TVAMMAPEVYIRLSLFKSKSDVMSFGVLLWELLTGEVYPRIDALAAVYANMKLTLP 325
QY 229 SSCPRFAEILHOCWEADAKKRSFKQIISLESMSNDTSLPDKCNSPLHNKAEKCEIE 288
DB 326 STPEPFARLLEECWPPDPHGRDPFGSILKRLVIEQSALFQMPLESFHSLOEDWLEIQ 385
QY 289 ATLERLKLKERDLSFKEOEL-----KERERLKNWEOKLTQOSNTPLLLPLAARMBE 341
DB 386 HMEDDLRTKEKELRSBEEELRLAAQORPOEBQLRRREDEIARENDIYERELHLMQDL 445
QY 342 SYFESTTESNSNEMSCQITATNSGEGKGNPSLOAMLMGFGDITSNKAKAVHESGNO 401
DB 446 SEKEPVRKRGKFNKESRLKLKREGGSHISLPS-----GFEH-----K 483
QY 402 INMQAKNSKSTKSRG 419

```

```

DB 484 ITVOA-----SPTLDKRGK 497

RESULT 2
M3K9 HUMAN STANDARD; PRT; 394 AA.
AC M3K9_HUMAN
ID M3K9_HUMAN
DT 01-FEB-1994 (rel. 28, Created)
DT 01-FEB-1994 (rel. 28, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 9 (EC 2.7.1.-) (Mixed
DE lineage kinase 1) (Fragment).
GN MAP3K9 OR MKK1 OR PRKEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; Pubmed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur J Biochem. 213:701-710(1993).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
PIR: S32467; J00229.
DR HSSP; P12931; 1FMK.
DR Genew; HGNC:6861; MAP3K9.
DR MIM; 600136; -.
DR GO; GO:0005524; F:ATP binding activity; NAS.
DR GO; GO:0004708; F:MAP kinase kinase activity; NAS.
DR GO; GO:0004668; P:protein amino acid phosphorylation; NAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR PRINTS; PR00109; TYRKINASE.
DR Pfam; PF00069; pkinaase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding.
FT NP_BIND 9 271 PROTEIN KINASE.
FT BINDING 30 17 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 ATP (BY SIMILARITY).
FT DOMAIN 127 127 BY SIMILARITY.
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;

Query Match 25.0%; Score 593; DB 1; Length 394;
Best Local Similarity 39.3%; Pred. No. 1.7e-31;
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

QY 15 DLQFFENCGGSGFSGVYRAKWSQDKYAVK-----KLKIEKAEILISVLSH 62
DB 2 EILTLEIIGIGFGKVFYAFWIGD--EVAVKARHDPDEDISQTIENVROEAKLPAMLKH 59
QY 63 RNIIOFYVILEPPVYGIYTERYASLSGLYDYNRSRSEMDMDHIMTATVDVAKGMHYLH 122
DB 60 PNIIALRGVCLKEPNLCLVMEFARGPLNRVLISGR--IPDIIIVNNAVQIARGNVYH 116
QY 123 MEAPVKYIHRDLKSRNVI-----AADGVLTICDGFASRFPHNTTHMSLVGTFPMMA 174

```

DB 117 DEAIPIIHRDLKSSNILLQKVENGDLSNKILKITDFTGLAREMHRRTTKSAAGTYAMA 176
 QY 175 PEVIOSLPSEICDTYSGVVMEMLTREVPFEGLEGLQVAMLYVEKNEELTPSSCPKS 234
 DB 177 PEVIRSMFSKSGSDVMSYGVLLWELLTGEVFFGIDGLRAYAVANNKUALPISCTCP 236
 QY 235 FAEILLQCEADAKKRPFKQIISILESMSNDTSLPDKNSFLHNKAEMCEIATELERL 294
 DB 237 FAKMEDCNPNPDHSPSPFNINILQDLTTEBSGFEMPKDSFICLODNKHEIQEMFDQ 296
 QY 295 KKLERDLSFEQEL-----KERERLKKMEQKLTQ 325
 DB 297 RAKEKELRTWEELTRALQCKQOEELLRREQLAER 334

RESULT 3
 M3KC_HUMAN STANDARD; PRT; 859 AA.
 ID M3KC_HUMAN STANDARD; PRT; 859 AA.
 AC 012852;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
 GN (leucine-zipper protein kinase) (ZPK).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCB1_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tetacarcinoma;
 RX MEDLINE=94311945; PubMed=8037767;
 RA Reddy U.R., Pleasure D.;
 RT "Cloning of a novel putative protein kinase having a leucine zipper domain from human brain.";
 RL Biochem. Biophys. Res. Commun. 202:613-620(1994).
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphotyrosine beta-casein, histone 1 and myelin basic protein in vitro.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol under basal conditions and dephosphorylated when membrane-associated (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U07358; AAA67343.1; -
 DR HSSP; P12931; 1FMK.
 DR Genew; HGNC:6851; MAP3K12.
 DR MIM; 600447; -
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:000586; C:plasma membrane; TAS.
 DR GO; GO:0007254; P:JNK cascade; TAS.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR0109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 125 366
 FT NP BIND 131 139
 FT BINDING 152 152
 FT ACT SITE 236 236
 FT DOMAIN 665 668
 FT POLY-PRO.
 FT POLY-GLU.
 SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;
 Query Match 23.8%; Score 565.5; DB 1; Length 859;
 Best Local Similarity 37.8%; Pred. No. 2.8e-29;
 Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;
 QY 10 QIKFDDLPFENCGGSGFSGSVYRAKWSODKEVAVKLLKIEKAEI--LSVLSHRNIIQ 67
 DB 119 EVDFEELIDLQWVGSAQAVFLGR--HGEVAVKVRDL-KETDIKHLRLKHNNIT 175
 QY 68 FPGVILEPPNYGIWVEYASLSLYDYINSRSEMDMHIMWATVAKGMYLMEAPY 127
 DB 176 FKGVCTQACYCICIMEFCAQGLYEVLARGPVTSPL--LVDSMGIAAGMYLHLH-- 230
 QY 128 KYIHRDLKSRNVVIAADGLKICDRGASR-FHNHTHMSLVGFPWMAPEVIOSLPVSET 186
 DB 231 KTIHRDLKSPNNILTYDDVAKISDFGTSKELSDKTSKMSFAGIVAMAPVLRNEPSEK 290
 QY 187 CDYSGVVLVEMLTREVPFEGLEGLQVAMLYVEKNERLTIPSSCPSPAEILLQCEAD 246
 DB 291 VDIWSFGVLLWELLTGEIYKVDSSAILTWGNSNLIHPVSSCGDGRKILLRQCMNSK 350
 QY 247 AKRPSFKQIISILESMSNDT-SLPDKNSFLHNKAEMCEIATELERL-----KIE 298
 DB 351 PNRRPSFQIILDLIASADVLSTPOE--TYPKSAEMREVEYVLEHEKIKSEGTCLHRL 408
 QY 299 RDLSPFKEG-----LKERERLKKMEQKLTQESNPLILLPLARMSSESYFSKTE 350
 DB 409 EELWMRRRELRLHALDIRH---YERKLERANN--LYVELNALMIQ---LEIKERE 456

RESULT 4
 M3KC_MOUSE STANDARD; PRT; 888 AA.
 ID M3KC_MOUSE STANDARD; PRT; 888 AA.
 AC Q60700; P70286;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing kinase) (DLK).
 GN MAP3K12 OR ZPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Brain;
 RX MEDLINE=95074107; PubMed=7983011;
 RA Holzman L.B., Merritt S.E., Fan G.;
 RT "Identification, molecular cloning, and characterization of dual leucine zipper bearing kinase. A novel serine/threonine protein kinase that defines a second subfamily of mixed lineage kinases.";
 RL J. Biol. Chem. 269:30808-30817(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
 RX MEDLINE=96365388; PubMed=8769565;
 RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
 RL DNA Cell Biol. 15:631-642(1996).
 RN [3]

```

RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX MEDLINE:96279259; PubMed:8663324;
RA Matsu M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
RT "Characterization of dual leucine zipper-bearing kinase, a mixed
RT lineage kinase present in synaptic terminals whose phosphorylation
RT state is regulated by membrane depolarization via calcineurin.";
RL J. Biol. Chem. 271:16888-16896(1996).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
CC testis, gastrointestinal tract, stomach, liver and pancreas.
CC Within the nervous system, predominantly expressed in neurons and
CC enriched in synaptic terminals.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14636; AAA57280.1; -.
DR EMBL: U23789; AAB17123.1; -.
DR PIR: A55318; A55318.
DR HSSP: P12931; 1FMK.
DR MGD: MGI:1346881; Map3K12.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50011; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN_KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP.
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
FT MUTAGEN 753 758 POLY-GLU.
FT MUTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
FT CONFILCT 192 192 E->A: NO CHANGE.
FT CONFILCT 18 18 V->A (IN REF. 2).
FT CONFILCT 28 29 K->N (IN REF. 2).
FT CONFILCT 382 382 S->T (IN REF. 2).
FT CONFILCT 494 495 EQ->D (IN REF. 2).
FT CONFILCT 517 517 N->D (IN REF. 2).
FT CONFILCT 794 794 E->G (IN REF. 2).
SQ SEQUENCE 888 AA; 96083 MW; CFECEFIJ34F889AB CRC64;
Query Match 23.8%; Score 565.5; DB 1; Length 888;
Best Local Similarity 37.8%; Pred. No.2.9e-29;
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;
QY 10 QIKFDDIQFFENCGGSGFSGVYRAKWIISQDEKVAVKKLKIEKEAEI--LSVLSHNNIIQ 67
DB 152 EVAFEEILDLQWQSGAGAVFLGRF--HGEBAVAVKVKNDL-KETIYIKHLKXKHPNIIIT 208

```

```

QY 68 FYGVILEPPNYGVITEVYASLSTLYDYNISNRSEEMDMHIMTATDVAKGMHYLHMAEVP 127
DB 209 FEGVCTQADPCYILMEFCAQGLVEVLRIGRVYTSLS--LVWMSGIAGSNVYLIH--- 263
QY 128 KYIHRDLKSRNVVIADVLKICDFGASR-FNNHTTHMSLVGTFPMAPEVIOQLPVSET 186
DB 264 KIHHDLSKPNMLIYDDVVKISDFGTSKELSDKSTKMSFATVAMMAPEVIRNEPVSER 323
QY 187 CPTYSYGVLMEMLTREVPFPGLEGLQVAMLVYKNERLTITSSCPSPFAELHQCWEAD 246
DB 324 VDIWSEFVGLWELTGEIPIYKDVDSALIWGVSNLSLPPVSSCPDGFKILLRCQMSK 383
QY 247 AKRPSFKIOIISLMSMSNDT-SLPDKNSPLHNAEMRCETEAETERLK-----KXE 298
DB 384 PNNRSPFRIHLHLDIASADVSTPOE--TYKSOQMRREVKLHFEKIKSGCTLHRL 441
QY 299 RDLSPKEQE-----LKERERLKMWEOKLTEOSNPLLLPLAARNSSESYFESKTEE 350
DB 442 EELVWRREELHMLADIREH-----YERKLERANN--LVWELNALMLQ--LELKERE 489

RESULT 5
M3XC RAT STANDARD; PRT; 888 AA.
AC 063756;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (MAPK-upstream kinase) (MUK).
GN MAP3K12 OR MUK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE:96226099; PubMed:8637721;
RX Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
RA "Activation of the JNK pathway by diabolically related protein kinases,
RT MEKK and MUK.";
RL Oncogene 12:641-650(1996).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D49785; BAA08621.1; -.
DR HSSP: P12931; 1FMK.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

```


DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 158 399
 FT NP_BIND 164 172
 FT BINDING 185 185
 FT ACT SITE 269 269
 FT DOMAIN 56 62
 FT DOMAIN 668 671
 FT DOMAIN 698 701
 FT DOMAIN 753 758
 SQ SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;
 Query Match 23.7%; Score 563.5; DB 1; Length 888;
 Best Local Similarity 37.8%; Pred. No. 3.9e-29;
 Matches 135; Conservative 70; Mismatches 117; Indels 35; Gaps 13;
 QY 10 QIKFDDLOFFENCGGSGFSGVYRAKWIISODKEVAVKKLKIEKEAEI--LSVLSHRNIIQ 67
 Db EVPFEEILDLQWVSGAGVFLGRF--HGEVAVKKVRL-KETDIKHLRKLNIIIT 208
 QY 68 FYGVILPEPPNYGIVTEVYASIGSLYDIYNSRSEMDMIMTAVDAKGMHYLNEAPV 127
 Db FKGVCTQAPCYCIIMEFCAGQGLYEVLRAGRPTPSD--LVDSMGAGGMNYLH-- 263
 QY 128 KVTHRLKSRNVYIADGVYKICDFGASR-FHNHTHMSLVGFPMAPREVLSQSLPVSET 186
 Db KTHRLKSRNVYIADGVYKICDFGASR-FHNHTHMSLVGFPMAPREVLSQSLPVSET 323
 QY 187 CDTSYGVVLMEMLTREVPFGLEGVAVLVVEKNERLTIPSSCRSPFAELHQCWEAD 246
 Db VDIWSGVVLMELTGEIPIKVDSSAIIWGVGNSLHLPVSSCGDGFILRLKQCNRK 383
 QY 247 AKRPSFKQIISLSEMSNDT-SLPDKNSFLNKAEMRCETATLERLK-----KLE 298
 Db PRNRPSPFQIILHLDAADVLSTPOE--TYFKSQAEWREVEVLAHEKIKSEGTCLRLLE 441
 QY 299 RDLSPFEEQ-----LKERERRLMQWQKLTQESNTPLPLPLARMBESYFEKTEE 350
 Db 442 EELWMRRERELRLHDLIRH---YERKLERANN--LYVELNALMLQ--LEIKERE 489
 RESULT 6
 CTR1_ARATH STANDARD; PRT; 821 AA.
 ID CTR1_ARATH STANDARD; PRT; 821 AA.
 AC 005609;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
 GN CTR1 OR AT5G03730 OR F17C15_150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsumo A., Muraki A., Nakayama S.,
 RA Nakasaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 Cell 72:427-441(1993).
 RT Arabidopsis, encodes a member of the raf family of protein kinases."
 RL Cell 72:427-441(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsumo A., Muraki A., Nakayama S.,
 RA Nakasaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 Cell 72:427-441(1993).

RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J., Becker M.,
 RA Stonking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozeresky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchhoff K., Toth K., King D., Bahner A., Miller B., Marra M.,
 RA Martensen R., McComble W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Slekema W., Pohl T.,
 RA Eutlian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,
 RA Ransberger U., Wedler E., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirks W., Moolman P., Klein Lankhorst R.,
 RA Weitenegger T., Botte G., Rose M., Hauf J., Benneiser S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gleten J., Ardiles W.,
 RA Bente O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana";
 RL Nature 408:823-826(2000).
 CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
 PATHWAY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
 CC -1- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
 RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
 COTYLEDON GROWTH IS IMPAIRED.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MIL/RAF SUBFAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L08789; AAA32779.1; -
 DR EMBL; L08790; AAA32780.1; -
 DR EMBL; AL162506; CAB82938.1; -
 DR PIR; T48400; T48400.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc1; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 65 69
 FT DOMAIN 135 141
 FT NP_BIND 551 809
 FT BINDING 557 565
 FT ACT SITE 578 578
 FT ACT SITE 676 676
 FT MUTAGEN 596 596
 FT MUTAGEN 694 694
 SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCDDCC15BC CRC64;
 Query Match 20.5%; Score 488; DB 1; Length 821;
 Best Local Similarity 40.7%; Pred. No. 2.8e-24;
 Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;
 QY 9 VOIKFDDLOFFENCGGSGFSGVYRAKWIISODKEVAVKKLKIEKEAEI--LSVLSHRNIIQ 67
 Db EVPFEEILDLQWVSGAGVFLGRF--HGEVAVKKVRL-KETDIKHLRKLNIIIT 208
 QY 68 FYGVILPEPPNYGIVTEVYASIGSLYDIYNSRSEMDMIMTAVDAKGMHYLNEAPV 127
 Db FKGVCTQAPCYCIIMEFCAGQGLYEVLRAGRPTPSD--LVDSMGAGGMNYLH-- 263
 QY 128 KVTHRLKSRNVYIADGVYKICDFGASR-FHNHTHMSLVGFPMAPREVLSQSLPVSET 186
 Db KTHRLKSRNVYIADGVYKICDFGASR-FHNHTHMSLVGFPMAPREVLSQSLPVSET 323
 QY 187 CDTSYGVVLMEMLTREVPFGLEGVAVLVVEKNERLTIPSSCRSPFAELHQCWEAD 246
 Db VDIWSGVVLMELTGEIPIKVDSSAIIWGVGNSLHLPVSSCGDGFILRLKQCNRK 383
 QY 247 AKRPSFKQIISLSEMSNDT-SLPDKNSFLNKAEMRCETATLERLK-----KLE 298
 Db PRNRPSPFQIILHLDAADVLSTPOE--TYFKSQAEWREVEVLAHEKIKSEGTCLRLLE 441
 QY 299 RDLSPFEEQ-----LKERERRLMQWQKLTQESNTPLPLPLARMBESYFEKTEE 350
 Db 442 EELWMRRERELRLHDLIRH---YERKLERANN--LYVELNALMLQ--LEIKERE 489
 RESULT 6
 CTR1_ARATH STANDARD; PRT; 821 AA.
 ID CTR1_ARATH STANDARD; PRT; 821 AA.
 AC 005609;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
 GN CTR1 OR AT5G03730 OR F17C15_150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsumo A., Muraki A., Nakayama S.,
 RA Nakasaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 Cell 72:427-441(1993).
 RT Arabidopsis, encodes a member of the raf family of protein kinases."
 RL Cell 72:427-441(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsumo A., Muraki A., Nakayama S.,
 RA Nakasaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 Cell 72:427-441(1993).

DB 544 MDIPWCDLNKIKKIGAGSFGTVHRAEMHSD--VAVKILMEDQFHAEVNEFLREVAINK 601
 QY 59 VLSHNNIOFYGVILEBPNYGVITEYASLGSLYDYN-SNRSEMDMDHMTWATVANG 117
 DB 602 RLHPHIVIFPMGAVTOPPNISIVTEYLSRGLYRLHKSARQDLERRRLSMAVYVANG 661
 QY 118 MYHLHMEAPVKVYHRDLKSRNVVIADGVLYKICDPGASRPHNNT--THMSLVCTPFMMAP 175
 DB 662 MNYLHNRNP-PIVHRDLKSPNLLVDKKYTKVCDPGLSRKASTPLSSKSAAGTPEMWAP 720
 QY 176 EVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVYKNERLTTPSCPRSE 235
 DB 721 EVLRDEPSNEKSVSESGVILMELATLQCPWGNLNPQVAAVGFCKRLIEPRNLPGV 780
 QY 236 AELLHQCWEADAKKRPSPFOIISILBSM 263
 DB 781 AALIEGCTWEPWKRPSFATIMDLRPL 808

RESULT 7
 M3K7_MOUSE STANDARD; PRT; 579 AA.
 ID M3K7_MOUSE
 AC 062073;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
 GN MAP3K7 OR TAK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP MEDLINE=96123277; PubMed=8531096;
 RA Yamauchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N., Taniguchi T., Nishida E., Matsunoto K.,
 RT Identification of a member of the MAPKKK family as a potential mediator of TGF-beta signal transduction.";
 RL Science 270:2008-2011(1995).
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
 CC -1- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annouce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D76446; BAA1184.1; --
 DR HSSP; P08631; IAD5.
 DR MGD; MGI:1346871; Map3K7.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000004; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 8 16 POLY-SER.
 FT DOMAIN 36 291 PROTEIN KINASE.
 FT NP_BIND 42 50 ATP (BY SIMILARITY).

FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 579 AA; 64227 MM; 97C8F6F3CE8233EE CRC64;
 Query Match 20.3%; Score 483; DB 1; Length 579;
 Best Local Similarity 30.2%; Pred. No. 3.7e-24;
 Matches 131; Conservative 82; Mismatches 151; Indels 70; Gaps 18;

QY 7 SEVQIKFDLDPFENCGGSGFSVYRAKISODKEVAVKLLIKIEAE-----IL 57
 DB 27 NEEEDIKYEIEVEVGRAPFVCKAKV--RAKVAIK--QISESEKKAFLVELROL 81
 QY 58 VLSHNNIOFYGVILEBPNYGVITEYASLGSLYDYN-SNRSEEM--DMDHMTWATD 114
 DB 82 SVYHNNIVKLYGACINP--VCLVMEYAGSLYVNLVH--AEPLPYIAAMAMKCLC 137
 QY 115 AKGMHYLHMEAPVKVYHRDLKSRNVVIADG-VLKICDPG-ASRPHNTHMSLVGTFW 172
 DB 138 SQGVAYLHSMQPKALIHRLDKPNLLVAGTVLKCIDPGTACDIQTHMTNKK--GSAAM 195
 QY 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE--LQVAMLVYKNERLTIPS 230
 DB 196 MAPEVEGSGSYSEKCDVFSWGIIWVIRRRKPFDEIGGPARIMW-AVHNGTRRPLIK 254
 QY 231 CRRSPAEHLHQCWEADAKKRPSPFOIISL-----ESMSNDT- 267
 DB 255 LRPPIESLNTKMSKDPGSRPMEELVKITLMRYFPADPELOYPCOYDEGGSNAT 314
 QY 268 ---SLPDKNSFLHNAKWERCE-IEATLERLKLERDLSFKQELKERRRRLKMEQKLT 323
 DB 315 STGSFMDIASNTNTSKSDPTNMEQVATNDTIRLSKLLKNQAKQSSGRLSLGASRGS 374
 QY 324 EGSNPLPLPLAARNSESEYFSKTEBSNSAMSCQITTSNGEGHGNPISQAMLMNF 383
 DB 375 SVESLP---PTS-----EGKMSADMSIEIRIVATA-GNQPRRRRSIODLVITGT 421
 QY 384 --GDIFSNMKAGAV 395
 DB 422 EPGQVSSRSSPSV 435

RESULT 8
 M3K7_HUMAN STANDARD; PRT; 606 AA.
 ID M3K7_HUMAN
 AC 043318; 043317; 043319;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
 GN MAP3K7 OR TAK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).
 RP TISSUE=Lung;
 RX MEDLINE=98153801; PubMed=9480845;
 RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
 RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism.";
 RL Biochem. Biophys. Res. Commun. 243:545-549(1998).
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
 CC -1- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1B;
 CC IsoId=O43318-1; Sequence=Displayed;
 CC Name=1A;
 CC IsoId=O43318-2; Sequence=VSP_004886;

```

CC      Name=1C;
CC      Ioid=04318-3; Sequence=VSP_004887, VSP_004888;
CC      SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP
CC      KINASE KINASE SUBFAMILY
CC      -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AB009357; BAA5026.1; -
DR      EMBL: AB009356; BAA5025.1; -
DR      EMBL: AB009358; BAA5027.2; -
DR      PIR: JCS955; JCS955.
DR      PIR: JCS956; JCS956.
DR      HSSP: P08631; IAD5.
DR      Genew: HGNC:6859; MAP3K7.
DR      MIM: 602614; -
DR      GO: GO:0004709; F-MAP kinase kinase kinase activity; TAS.
DR      GO: GO:0007179; P-TGFPbeta receptor signaling pathway; TAS.
DR      InterPro: IPR000719; Prot kinase.
DR      InterPro: IPR002230; Ser Thr kinase.
DR      InterPro: IPR001245; Tyr_kinase.
DR      Pfam: PF00069; Kinase; 1.
DR      PRINTS: PR00109; TYRKINASE.
DR      ProDom: PD000001; Prot kinase; 1.
DR      PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR      PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR      PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR      Transfaser: Serine/threonine-protein kinase; ATP-binding;
KW      Alternative splicing.
FT      DOMAIN 8 14 POLY-SER.
FT      NP_BIND 36 291 PROTEIN KINASE.
FT      BINDING 42 50 ATP (BY SIMILARITY).
FT      ACT_SITE 156 63 ATP (BY SIMILARITY).
FT      VARSP_LIC 404 430 BY SIMILARITY.
FT      VARSP_LIC 509 518 Missing (in isoform 1A).
FT      VARSP_LIC 518 /FTid=VSP_004886.
FT      VARSP_LIC 519 /FTid=VSP_004887.
FT      VARSP_LIC 519 /FTid=VSP_004888.
FT      VARSP_LIC 519 Missing (in isoform 1C).
FT      SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;
Query Match 19.9%; Score 471.5; DB 1; Length 606;
Best Local Similarity 29.7%; Pred. No. 2.2e-23;
Matches 133; Conservative 85; Mismatches 151; Indels 79; Gaps 20;
QY      7 SFVQIKFDLPQFENCGGSGFSGVSAKYAKISQDEYAVVKL-LKIEKEAI-----LSVL 60
DB      27 NFEIIDEYKEIEVEVGRGAFVVCCKAM-RAKDVAIKQIESESRKAFIVELRLSLRY 84
QY      61 SRHNIIQFQVILPEPPNYGIVTYASLGSIVYINSRSEEM--DMDHIMTATVANG 117
DB      85 NHENIYKLTGACINP--VCLVMEYAEAGGSLYNVLAHG--AEPLPYTAALHAMSWCLQCSQG 140
QY      118 MHYLFHEAPKVIHRDLKSRNVIAADG-VLAKICDGG-ASRFNHTHTMSLVCTFFPMAP 175
DB      141 VAYLHSMQPKALHRLDKPNNLLVAGGYVLKICDGTACDICTHTMTNKN--GSAAMMAP 198
QY      176 EVIQLSPVSTCDTYSYGVVLMELTRVPEFKLEG--LQVAMLVKKERLTISSCR 233
DB      199 EVEEGSNYSBKCVFSGWGLTMEVITRRKFEIDGGAPRIMV-AVHNGRRPPLINLPR 257
QY      234 SPFELHQCHEADAKKRPFKQIISL-----ESMSNDT----- 267
DB      258 PLESIMTRCWSKDPSPSPMEELVKIMTHLMRYPGADEPLQYPCOYSEGGNSATSTG 317
QY      268 SLDPDKNSFLHNKAEMRC-IEATVLERLKLERDLSPKEGELMERERRLKMWQKULTGEG 326

```

```

DB      318 SMDIASITNTSKSDPTNNEQVATNDTIKRLSKLLKQAKQSSGRLSLGASRESSVE 377
QY      327 NPPLLPLPLAARMSESYESTKTEESNSAEMSCQITATS---NGEGHNPBLQAMLMG 382
DB      378 SLP---PTS-----EGKEMSADMSEIEARIATVYSKPKGHRKTAS----- 417
QY      383 PGDIFSMNAGAVMHSQGINNOAKONS 410
DB      418 FGNILDVPE--TVISG--NGQPRRS 439
RESULT 9
ABL_MTVAB
ID      ABL_MTVAB STANDARD; PRT; 746 AA.
AC      P00521;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).
GN      V-ABL.
OS      Abelson murine leukemia virus.
OC      Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX      NCBI_TaxId=11788;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=83221648; PubMed=6304726;
RA      Reddy E.P., Smith M.J., Srinivasan A.;
RT      "Nucleotide sequence of Abelson murine leukemia virus genome:
RT      structural similarity of its transforming gene product to other onc
RT      gene products with tyrosine-specific kinase activity.";
RL      Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
RN      [2]
RP      REVISIONS TO 588-746.
RA      Reddy E.P., Smith M.J., Srinivasan A.;
RL      Proc. Natl. Acad. Sci. U.S.A. 80:7372-7372(1983).
RN      [3]
RP      SEQUENCE OF 233-327 FROM N.A.
RX      MEDLINE=83245023; PubMed=6191223;
RA      Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
RT      "Homology between phosphotyrosine acceptor site of human c-abl and
RT      viral oncogene products.";
RL      Nature 304:167-169(1983).
CC      -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC      tyrosine phosphate.
CC      -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
CC      POLYPROTEIN.
CC      -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC      SUBFAMILY.
CC      -1- SIMILARITY: Contains 1 SH2 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: VO1541; -; NOT ANNOTATED_CDS.
DR      EMBL: K00010; AAA6470.1; -
DR      HSSP: P00519; 2ABU.
DR      InterPro: IPR000719; Prot_kinase.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001245; Tyr_kinase.
DR      Pfam: PF00069; Kinase; 1.
DR      Pfam: PF00017; SH2; 1.
DR      PRINTS: PR00401; SH2DOMAIN.
DR      PRINTS: PR00109; TYRKINASE.
DR      ProDom: PD000001; Prot_kinase; 1.
DR      ProDom: PD000093; SH2; 1.
DR      SMART: SM00252; SH2; 1.
DR      SMART: SM00219; Tyrc; 1.

```

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00001; SH2; 1.
 KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;
 KM SH2 domain.
 FT DOMAIN 13 103 SH2.
 FT DOMAIN 128 379 PROTEIN KINASE.
 SQ SEQUENCE 746 AA; 81872 MW; B9072FF5FE9257 CRC64;
 Query Match 18.3%; Score 434; DB 1; Length 746;
 Best Local Similarity 28.7%; Pred. No. 7.7e-21;
 Matches 121; Conservative 73; Mismatches 172; Indels 56; Gaps 13;
 QY 15 DLQFFNCGGSGSYRAKVIQSDKEVANKL-----LKE---KKAELISVLSHNITQ 67
 DB 127 DITMKHKLGGQYGEVYEGWKKYSLTVAVKTLKEDTMEVEEFLKEAAVKKELKHNLVQ 186
 QY 68 FVGVIIEPPNYGIVTVYASISLYDIYNSRSEMDIMTWATDVAKGMHYLMEAPV 127
 DB 187 LAGVCRREPFFYITTFMTYGNLDDYLRGNROGVANVILLNATQISAMEYLEKK--- 243
 QY 128 KVIRHDLKSRNVIAADGVLIKIDFGASRPHNHTHSLVGT--P--WMAPEVIQSLPV 184
 DB 244 NFIRHDLAARNCLVGNHLVKVADFGLSRLMTGDTYTAHAGKFPKMTAPESLAYNKFS 303
 QY 185 ETCDTTSYGVVLEMLTREV--PKKLEGLQVAVLVKNERLTPSSCRSFALLHQQW 243
 DB 304 IKSDVAFGLVLEIATYGMSPPGIDLSGV--YELLEKDRMRPGCPEKYEYELMRACW 362
 QY 244 EADAKRPFKQIISLESMSNDTSLPDKNSFLHNKAERCEIEATLETKLERD--- 300
 DB 363 QMPSRPRPFAELHQAFTMFGSSISD-----EVEKELG--KGTGKGAG 406
 QY 301 --LSFKQELKEBERRLKWEQQLTEQSNTPLLPLAARNSESYESKTESNSAEMSC 358
 DB 407 SMLQADELPKTKR--RAAEQKASPPSLPKLLRQVTAPSPSGLSHKKEATKGSAGM 465
 QY 359 QITATNGEGHGMNPSLQAMLMFGFDIFSMNKAQAVMISGMQINNOAKNSKITSKR 418
 DB 466 GTPATA-----EPAPFS-----NRYGLSKASSEEMRVRRHNGSESPGRDK 506
 QY 419 GK 420
 DB 507 GR 508
 RESULT 10
 ABL1 HUMAN STANDARD; PRT; 1130 AA.
 ID ABL1 HUMAN Q13869; Q13870; Q16133;
 AC P00519; Q13869 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150) (c-ABL).
 GN ABL1 OR ABL OR JTK7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibroblast;
 RA MEDLINE=90082420; PubMed=2687768;
 RA Fainstein E., Einat M., Gokkel E., Marcelle C., Croce C.M.,
 RA Gale R.P., Canaan E.;
 RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs.";
 RL Oncogene 4:1477-1481 (1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RA MEDLINE=87028219; PubMed=3021337;
 RA Shivelman E., Lifshitz B., Gale R.P., Roe B.A., Canaan E.;

RT "Alternative splicing of RNAs transcribed from the human abl gene and
 RT from the bcr-abl fused gene.";
 RL Cell 47:277-284 (1986).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS IA AND IB).
 RC TISSUE=Lung carcinoma;
 RX MEDLINE=95394474; PubMed=7665185;
 RA Chisoe S.L., Bodenreich A., Wang Y.-F., Wang Y.-P., Burian D.,
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,
 RA Melauri H.-U., Pan H.-Q., Sathian O.H., Toch S., Wang Z., Zhang G.,
 RA Heisterkamp N., Groffen J., Roe B.A.;
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and
 RT regions involved in the Philadelphia chromosome translocation.";
 RL Genomics 27:67-82 (1995).
 RN [4]
 RP SEQUENCE OF 360-426 FROM N.A.
 RX MEDLINE=83245023; PubMed=6191223;
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
 RT "Homology between phosphotyrosine acceptor site of human c-abl and
 RT viral oncogene products.";
 RL Nature 304:167-169 (1983).
 RN [5]
 RP SEQUENCE OF 27-40 FROM N.A.
 RX MEDLINE=88065859; PubMed=2825022;
 RA Fainstein B., Marcelle C., Rosner A., Canaan E., Gale R.P.,
 RA Drezner O., Smith S.D., Croce C.M.;
 RT "A new fused transcript in Philadelphia chromosome positive acute
 RT lymphocytic leukaemia.";
 RL Nature 330:386-388 (1987).
 RN [6]
 RP SEQUENCE OF 825-845 FROM N.A.
 RX MEDLINE=94142331; PubMed=7545908;
 RA Inokuchi K., Futaki M., Dan K., Nomura T.;
 RT "Sequence analysis of the mutation at codon 834 and the sequence
 RT variation of codon 837 of c-abl gene.";
 RL Leukemia 8:343-344 (1994).
 RN [7]
 RP STRUCTURE BY NMR OF SH2 DOMAIN.
 RX MEDLINE=92370689; PubMed=1505033;
 RA Overduin M., Rios C.B., Mayer B.J., Baltimore D., Cowburn D.;
 RT "Three-dimensional solution structure of the src homology 2 domain of
 RT c-abl.";
 RL Cell 70:697-704 (1992).
 RN [8]
 RP STRUCTURE BY NMR OF SH2 DOMAIN.
 RX MEDLINE=93101588; PubMed=1281542;
 RA Overduin M., Mayer B.J., Rios C.B., Baltimore D., Cowburn D.;
 RT "Secondary structure of Src homology 2 domain of c-Abl by
 RT heteronuclear NMR spectroscopy in solution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677 (1992).
 RN [9]
 RP STRUCTURE BY NMR OF SH3 DOMAIN.
 RX MEDLINE=96318178; PubMed=8590002;
 RA Gosses Y.Q., Zheng J., Overduin M., Mayer B.J., Cowburn D.;
 RT "The solution structure of Abl SH3, and its relationship to SH2 in
 RT the SH(32) construct.";
 RL Structure 3:1075-1086 (1995).
 RN [10]
 RP 3D-STRUCTURE MODELING OF SH3 DOMAIN.
 RX MEDLINE=95199229; PubMed=7892170;
 RA Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;
 RT "Homology modeling of the Abl-SH3 domain.";
 RL Proteins 20:203-215 (1994).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.
 RX MEDLINE=96398698; PubMed=8805596;
 RA Nam H.-J., Haeser W.G., Roberts T.M., Frederick C.A.;
 RT "Intramolecular interactions of the regulatory domains of the Bcr-Abl
 RT kinase reveal a novel control mechanism.";
 RL Structure 4:1105-1114 (1996).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.
 RX MEDLINE=98365516; PubMed=9698566;

RA Pisabarro M.T., Serrano L., Wilmanns M.;
 RT "Crystal structure of the abl-SH3 domain complexed with a designed
 RT high-affinity peptide ligand: implications for SH3-ligand
 RT interactions." J. Mol. Biol. 281:513-521(1998).
 RN (13)
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 223-515.
 RX MEDLINE=20446271; PubMed=10988075;
 RA Schindler T., Bornmann W., Pellicena P., Miller W.T., Clarkson B.,
 RA Kurtyan J.;
 RT "Structural mechanism for STI-571 inhibition of abelson tyrosine
 RT kinase." Science 289:1938-1942(2000).
 RL Science 289:1938-1942(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=IA;
 CC IsoId=P00519-1; Sequence=Displayed;
 CC Name=IB;
 CC IsoId=P00519-2; Sequence=VSP_004957;
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- DISEASE: PARTICIPATES IN A T(9;22)(Q34;Q11) CHROMOSOMAL
 CC TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR
 CC CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND
 CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chromosome/Genes/ABL.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL; X16416; CAA34438.1; -;
 DR EMBL; M14752; AAA51561.1; -;
 DR EMBL; U07563; AAB60394.1; -;
 DR EMBL; U07563; AAB60393.1; -;
 DR EMBL; U07561; AAB60393.1; JOINED.
 DR EMBL; S69223; AAD14034.1; -;
 DR PIR; S08519; TVHUA.
 DR PDB; 1AB2; 31-JAN-94.
 DR PDB; 2AB2; 04-SEP-97.
 DR PDB; 1AWO; 28-JAN-98.
 DR PDB; 1BB2; 25-NOV-98.
 DR PDB; 1FPU; 20-SEP-00.
 DR PDB; 1ABL; 01-NOV-94.
 DR PDB; 1JUS; 06-NOV-02.
 DR Genew; HGNC:76; ABL1.
 DR MIM; 189980; -;
 DR GO; GO:0004713; F:protein tyrosine kinase activity; TAS.
 DR GO; GO:0006330; P:induction of apoptosis by DNA damage; TAS.
 DR GO; GO:0006398; P:ismatch repair; TAS.
 DR GO; GO:0007048; P:oncogenesis; TAS.
 DR GO; GO:0000704; P:regulation of cell cycle; TAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR GO; GO:0001115; P:s-specific transcription in mitotic cell cycle; TAS.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR PRINTS; PR00401; SH2DOMAIN.

DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR ProDom; PD000093; SH2; 1.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00219; TYRKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 1.
 KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
 KW Phosphorylation; SH2 domain; SH3 domain; Chromosomal translocation;
 KW 3D-structure; Alternative splicing.
 FT DOMAIN 61 121
 FT DOMAIN 127 217
 FT DOMAIN 242 493
 FT DOMAIN 605 609
 FT DOMAIN 18 22
 FT DOMAIN 605 609
 FT DOMAIN 782 1019
 FT DOMAIN 897 903
 FT SITE 26 27
 FT NP_BIND 248 256
 FT BINDING 271 271
 FT ACT_SITE 363 363
 FT MOD_RES 393 393
 Query Match 17.3%; Score 410; DB 1; Length 1130;
 Best Local Similarity 25.9%; Pred. No. 4,7e-19;
 Matches 119; Conservative 80; Mismatches 183; Indels 78; Gaps 15;
 QY 15 DLQFFNCGGSGFSGSYRAKWIISODEVAVKLU---LKIR---KEAEILSVSHNIIQ 67
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 300
 QY 241 DITMKHKLGGGGGYEGYEGWKKYSLTAVKTLKEDTMEVEEFLEKAAVKKETKHNVLQ 300
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 357
 QY 68 FGVILPEPPNYGIVTEYASLGLYDYINSREEMDMHMTWATDVAKGMHLYHEAPV 127
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 357
 QY 301 LIGVCTREPPPYITIFPMYTGNDLDFRCNQEVAAVLLVWATQISSAMEYLEKK--- 357
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 417
 QY 128 KYIHRDLKRNVIYAADVLKICDRGASRFNHTTMSLVGR-FP--WMAPEYIGSLPVS 184
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 417
 QY 358 NFTHRLAARNCVLGEMNLVAVADFGLSRLMTGDTYTAAGAKFPLKWAPESLAVNKS 417
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 476
 QY 185 ETCDTSYGVVLMEMLTREV-PFKGLEGLQVAMLVVEKNERLTIPSSCRSFAELHQCV 243
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 476
 QY 418 IKSVDWAFGLVMEITTYGMSYPFGIDLSQV-YELLEKDYRMRERPGCEKEYLEMRACH 476
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 525
 QY 244 EADAKKRPSPKOIISLESMSNDTSLPDKCSFNLHKAEMRCIEATLRLKKERDLSE 303
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 525
 QY 477 QNPPSDRPSFAFHQFETMPQESSISDEVEKELQGG-----VRGAVSTL-----LQA 525
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 570
 QY 304 KEQELKERERRLKMEQKLTBOSNTPLLLPLAARSESEYFPSKTE-ESNSAE---MGC 358
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 570
 QY 526 PELPFTKTRTSRRAAEHRDTTDPVEMP-----HSKGGESDPLDHEPAVSP 570
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 613
 QY 359 QITATNGSGHGMNPQLQAMLMFGFDISNMKAGAVHSGMGINN-----QAKNSKKT 413
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 613
 QY 571 LIPKRRGPPGEG-----GLNEDERLRLPKKKYTNLPSALLIKKKKTAFT 613
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 643
 QY 414 TSKRRGKKVNMALGSPDPLSGDDDDDDGEEEDNDMN 453
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 643
 QY 614 PPKRS-----SSFREMGGQPPRRGAGEEGRDIN 643
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 643
 RESULT 11
 ABL1_MOUSE STANDARD; PRT; 1123 AA.
 AC P00520; P97896; Q61252; Q61254; Q61255; Q61256; Q61257;
 AC Q61258; Q61259; Q61260; Q61261;

DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150
DE (c-ABL).
GN ABL1 OR ABL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=88068561; PubMed=3317402;
RA Oppi C., Shore S.K., Reddy E.P.;
RT "Nucleotide sequence of testis-derived c-abl cDNAs: implications for
RT testis-specific transcription and abl oncogene activation.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8200-8204(1987).
RN
RN [2]
RP SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I; II; III AND IV).
RX MEDLINE=95394474; PubMed=766185;
RA Chisoso S.L., Bobentich A., Wang Y.-F., Butian D.,
RA Clifton S.W., Cabretree J., Freeman A., Iyer K., Jian L., Ma Y.,
RA McNaury H.-U., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,
RA Heisterkamp N., Gelfand J., Roe B.A.;
RT "Sequence and analysis of the human ABL gene, the BCR gene, and
RT regions involved in the Philadelphia chromosomal translocation.";
RL Genomics 27:67-82(1995).
RN
RN [3]
RP SEQUENCE OF 85-182 FROM N.A.
RX MEDLINE=84106840; PubMed=6319018;
RA Wang Y.-Y., Ledley F., Goff S., Lee R., Groner Y., Baltimore D.;
RT "The mouse c-abl locus: molecular cloning and characterization.";
RL Cell 36:349-356(1984).
RN
RN [4]
RP ALTERNATIVE SPLICING.
RX MEDLINE=88202920; PubMed=3283651;
RA Bernards A., Paskind M., Baltimore D.;
RT "Four murine c-abl mRNAs arise by usage of two transcriptional
RT promoters and alternative splicing.";
RL Oncogene 2:297-304(1988).
RN
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 61-121.
RX MEDLINE=95393198; PubMed=7664083;
RA Musacchio A., Saraste M., Williams M.;
RT "High-resolution crystal structures of tyrosine kinase SH3 domains
RT complexed with proline-rich peptides.";
RL Nat. Struct. Biol. 1:546-551(1994).
CC
CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC
CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC; THE MYRISTOYLATED C-ABL PROTEIN
CC WAS REPORTED TO BE NUCLEAR.
CC
CC -1 ALTERNATIVE PRODUCTS:
CC Name=i;
CC IsoId=P00520-1; Sequence=Displayed;
CC Name=ii;
CC IsoId=P00520-2; Sequence=VSP_004959;
CC Name=iii;
CC IsoId=P00520-3; Sequence=VSP_004958;
CC Name=iv;
CC IsoId=P00520-4; Sequence=VSP_004960;
CC -1 TISSUE SPECIFICITY: Widely expressed.
CC -1 SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1 SIMILARITY: Contains 1 SH2 domain.
CC -1 SIMILARITY: Contains 1 SH3 domain.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J02995, AAA88241.1, -.
DR EMBL, U14721, AAB60451.1, -.
DR EMBL, U14720, AAB60451.1, JOINED.
DR EMBL, U14721, AAB60450.1, -.
DR EMBL, U14720, AAB60450.1, JOINED.
DR EMBL, U14721, AAB60448.1, -.
DR EMBL, U13835, AAB60448.1, JOINED.
DR EMBL, U14721, AAB60449.1, -.
DR EMBL, U13835, AAB60449.1, JOINED.
DR EMBL, X07539, CA30411.1, -.
DR EMBL, X07540, CA30413.1, -.
DR EMBL, X07541, CA30414.1, -.
DR EMBL, M12263, AAA37136.1, -.
DR EMBL, M12264, AAA37137.1, -.
DR EMBL, M12265, AAA37138.1, -.
DR EMBL, M13266, AAA37134.1, -.
DR EMBL, K03228, AAA37135.1, -.
DR PDB, IABQ, 15-OCT-95.
DR PDB, IABQ, 15-OCT-95.
DR PDB, IEPY, 20-SEP-00.
DR PDB, IIEP, 18-APR-01.
DR PDB, IME2, 18-SEP-02.
DR MGD, MGI:87859, Ab1.1.
DR InterPro, IPR000719, Prot_kinase.
DR InterPro, IPR000980, SH2.
DR InterPro, IPR001452, SH3.
DR InterPro, IPR001245, Tyr_kinase.
DR Pfam, PF00069, Dkinase, 1.
DR Pfam, PF00017, SH2, 1.
DR Pfam, PF00018, SH3, 1.
DR PRINTS, PRO0401, SH2DOMAIN.
DR PRINTS, PRO0452, SH3DOMAIN.
DR PRINTS, PRO0109, TYRKINASE.
DR ProDom, PD000001, Prot_kinase, 1.
DR ProDom, PD000093, SH2, 1.
DR ProDom, PD000066, SH3, 1.
DR PROSITE, PS00107, PROTEIN_KINASE_ATP, 1.
DR PROSITE, PS00109, PROTEIN_KINASE_TYR, 1.
DR PROSITE, PS50011, PROTEIN_KINASE_DOM, 1.
DR PROSITE, PS50001, SH2, 1.
DR PROSITE, PS50002, SH3, 1.
DR Transferrase, Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
KW Phosphorylation; SH2 domain; SH3 domain; Alternative splicing;
KW Nuclear protein; 3d-structure.
FT DOMAIN 61 121 SH3.
FT DOMAIN 127 217 SH2.
FT DOMAIN 242 493 PROTEIN KINASE.
FT DOMAIN 605 609 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 18 22 POLY-SER.
FT DOMAIN 605 609 POLY-LYS.
FT DOMAIN 804 1012 PRO-RICH.
FT DOMAIN 891 897 POLY-PRO.
FT NP_BIND 248 256 ATP (BY SIMILARITY).
FT BINDING 271 271 ATP (BY SIMILARITY).
FT ACT_SITE 363 363 BY SIMILARITY.
FT MOD_RES 393 393 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSSSCYCLE -> MSQRTYTKCR
FT VARSPPLIC 1 26 VQRPALPFW (in isoform III).
FT VARSPPLIC 1 26 /FtId=VSP_004958.
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSSSCYCLE -> MISFDLSDEL
FT VARSPPLIC 1 26 HIKLVLDV (in isoform II).
FT VARSPPLIC 1 26 /FtId=VSP_004959.
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSSSCYCLE -> MGQCPGKVLGD
FT VARSPPLIC 1 26 QRRSLPALHFIKAGKGRDSSRHGGPHCNVEEH (in
FT VARSPPLIC 1 26 isoform IV).
FT VARSPPLIC 1 26 /FtId=VSP_004960.
FT VARSPPLIC 1 26 LYS -> VGDW (IN REF. 2).

```

```

FT TURN 77 78
FT STRAND 79 79
FT STRAND 82 82
FT TURN 84 85
FT STRAND 87 93
FT TURN 95 96
FT STRAND 99 104
FT TURN 105 106
FT STRAND 107 112
FT HELIX 113 115
FT STRAND 116 118
SQ SEQUENCE 1123 AA; 122676 MW; 284F0830644AFD8F CRC64;

Query Match 17.1%; Score 405; DB 1; Length 1123;
Best Local Similarity 27.2%; Pred. No. 9,9e-19;
Matches 125; Conservative 78; Mismatches 179; Indels 78; Gaps 17;

15 DLQFENCGSGFGSVYRAKWTISODKEVAVKKL---LKE--KEAEILSVLSHRNIIQ 67
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 DITMKHKKGQGYGEVYEGVWKYSLTVAVKLKEPTMEVEELKKAAMKEIKHNVLVQ 300
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
68 FVGVLEPPNYGIVTEYASLGSLVDYINSRSEEMDMHMTWADVDVAKMYLHMEAPV 127
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 LIGVCTREPPFYIIEFMTYGNLIDYLRECNQEVSAVILLVWATQISSAMEYLEKK--- 357
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
128 KVIHRDLKSRNVVIADVGLKICDFGASRFNHTTMSLVGT-PP--WMAPEYIOQLPVS 184
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
358 NFIHRDLARNCLVGNHLYKADPGLSLMTGDIYTAAGAKFPIKWTAPESLAYNKFS 417
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 ETCDITYSYGVVLEMTREV-PEKLEGLQVAVLVKERNRLTIPSCPSFAELLHQCW 243
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
418 IKSDWAFQVFLMEITTYGMSPPGIDLSQV-YELLEKQYRMERPGCEKEVYELMRACW 476
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 EADAKRPSFKQIISLSEMSNDTSLPDKNSFLNKAEMRCEIETALREKLKLEED--- 300
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
477 QMPSRPSFAEIHQAFETMFOESSISD-----EVEKELG--KRGTRGAG 520
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 --LSFEGQLKERERLKMWEQKLTQOSNTPLPLPLAARMSESYESTKEENSGAEMGC 358
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
521 SMLQAPBLPTKTRTC-RAEQK--DAPITPELL-----HTKGESALDSEPA-VSP 570
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 QTTATNSGEGHGNPFLQAMMLMGFDIFSNNKAGAVHSGMOINM-----QAKONSSKT 413
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
571 LLPKRGKPPDG-----SLNEDDELPRDKRTNLPALIKKKKKMAFT 613
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
414 TSKRKGKVMALGFSDFPLSGDDDDDDGEEEDNDMN 453
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
614 PKRS-----SSFREMDOGPRRGASEDDSDRELGN 643
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
ABL2 HUMAN STANDARD; PRT; 1182 AA.
ID ABL2 HUMAN
AC P42684;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase ABL2 (EC 2.7.1.112) (Tyrosine kinase Arg).
GN ABL2 OR ARG OR ABL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORMS IA AND IB).
RX MEDLINE=90332670; PubMed=2198571;
RA Krub G.D., Perego R., Mikl T., Aaronson S.A.;
RT "The complete coding sequence of arg defines the ABLson subfamily of
RT cytoplasmic tyrosine kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5802-5806 (1990).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

```

```

CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IB;
CC IsoId=P42684-1; Sequence=Displayed;
CC Name=IA;
CC IsoId=P42684-2; Sequence=VSP 004961;
CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- DATABASE: NAME=AbLae Genet. CytoGene. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/ABL2ID226.html".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M35296; AAA35553.1; -
CC PIR; A35962; A35962.
CC PIR; B35962; B35962.
CC HSSP; P00519; 1BBZ.
CC GeneW; HGNC:77; ABL2.
CC MIM; 164690; -
CC InterPro; IPR000719; Prot_Kinase.
CC InterPro; IPR000980; SH2_
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00452; SH3DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_Kinase; 1.
CC ProDom; PD000093; SH2; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC SMART; SM00219; TYRK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00001; SH2; 1.
CC PROSITE; PS00002; SH3; 1.
CC KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
CC Phosphorylation; SH2 domain; SH3 domain; Alternative splicing.
CC FT DOMAIN 107 167
CC FT DOMAIN 173 263
CC FT DOMAIN 288 539
CC FT DOMAIN 561 564
CC FT DOMAIN 658 660
CC FT DOMAIN 732 739
CC FT DOMAIN 843 1055
CC FT DOMAIN 984 988
CC FT NP BIND 294 302
CC FT BINDING 317 317
CC FT ACT_SITE 409 409
CC FT MOD_RES 439 439
CC FT VARSPPLIC 1 73
CC FT
CC SQ SEQUENCE 1182 AA; 128343 MW; ED93869BC2B14FMA CRC64;
CC /FTID=VSP_004961.
Query Match 17.1%; Score 405; DB 1; Length 1182;
Best Local Similarity 32.8%; Pred. No. 1,1e-18;
Matches 113; Conservative 59; Mismatches 144; Indels 28; Gaps 13;

```

```

QY 5 GASFVQIKFD---DIQFENCGSGFSGSYRAKWTISQKEVAVKTL-----LKIE---KE 53
DB 273 GVSPIHDKMEKMTDTIMTKHKLREGGQIGEVYGVWKKYSILTAVAKTLKEDTMEVEEPLKE 332
QY 54 AELLISVLSHRNIIQFYGV-ILBPPNYGIVTEVASLSGLYDINSNREEMDMHMTWAT 112
DB 333 AAVMKELIKHPNLVQLGVCITLLEPPY-IVTEWMPYCNLDLYLNCRREETAVALLMYMT 391
QY 113 DVAKGHHYLMEMPVKYIHRDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGT-PP 171
DB 392 QISSNAEYLEKK---NFIHRDLAARNCLVGENHVAVADFGSLRMTGDTYTAHAGAKEP 448
QY 172 --WMAPEVIGLSLVSENCDTYSYGVLMEMLTREV-PEKGLGLOVAMLVENKNERLTIP 228
DB 449 IKMTAESLALYNTFTSIKSDVMAWGLVLMELATIGMSDYPETIDSGV-YDLLEKGYRMEQD 507
QY 229 SSCPRSEAFELHOCWEADAKKRPSPFOIISILEMSNDTSLPDKCNSFLHNKAEMRCEIE 288
DB 508 ECGPPRYVELMRACWKMSPADRSFPAETHQAFETMHDSSISEVAEEL-GRASSSSSV 566
QY 289 ATERLKLERDLSFKEQELKEKERRLKWE--OKLTQSGNTPL 330
DB 567 PYLPRPLPLPS---KTRTLKKQVENKENIEGAQADATENSASSL 606

RESULT 13
ABL_DROME STANDARD; PRT; 1520 AA.
ID ABL_DROME
AC P00522;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-ash).
GN ABL OR DASH OR ABL-1.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=88174728; PubMed=2832740;
RA Henkemeyer M.J., Bennett R.L., Gertler F.B., Hoffmann F.M.;
RT "DNA sequence, structure, and tyrosine kinase activity of the
RT Drosophila melanogaster Abelson proto-oncogene homolog.";
RL Mol. Cell. Biol. 8:843-853(1988).
RN 2
RP SEQUENCE OF 374-648 FROM N.A.
RX MEDLINE=84082064; PubMed=6317185;
RA Hoffmann F.M., Fresco L.D., Hoffmann-Falk H., Shilo B.-Z.;
RT "Nucleotide sequences of the Drosophila src and abl homologs:
RT conservation and variability in the src family oncogenes.";
RL Cell 35:393-401(1983).
RN 3
RP FUNCTION.
RX MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Peller M.;
RT "Roles of Amsadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
CC -1- FUNCTION: ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS
CC -1- JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19692; AAA28934.1; -.
DR EMBL; M19690; AAA28934.1; JOINED.
DR EMBL; M19691; AAA28934.1; JOINED.
DR EMBL; K01042; AAA28443.1; -.
DR PIR; A28128; TVFA.
DR HSSP; P00519; IAB2.
DR FLYBase; FBgn0000017; Abl.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR GO; GO:0005927; C:muscle tendon junction; IDA.
DR GO; GO:0004713; F:protein tyrosine kinase activity; IDA.
DR GO; GO:0007411; P:axon guidance; IMP.
DR GO; GO:0007417; P:central nervous system development; IGI.
DR GO; GO:0007391; P:dorsal closure; NAS.
DR GO; GO:0002009; P:morphogenesis of an epithelium; NAS.
DR GO; GO:0008360; P:regulation of cell shape; NAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2_kinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
KW transferase; tyrosine-protein kinase; ATP-binding; phosphorylation;
KW SH2 domain; SH3 domain.
FT DOMAIN 204 265
FT DOMAIN 271 363
FT DOMAIN 388 644
FT NP BIND 394 402
FT BINDING 417 417
FT ACT_SITE 509 509
FT MOD_RES 539 539
FT CONFLICT 374 377
FT CONFLICT 645 648
SQ SEQUENCE 1520 AA; 161836 MW; AD6A5060579EAD7B CRC64;

Query Match 16.9%; Score 402; DB 1; Length 1520;
Best Local Similarity 26.3%; Pred. No. 2,3e-18;
Matches 131; Conservative 69; Mismatches 182; Indels 116; Gaps 17;

```

```

QY 15 DIQFENCGSGFSGSYRAKWTISQKEVAVKTL-----LK-IEKAEILSVLSHRNIIQ 67
DB 387 DTMKHKLGSGGQIGEVYGVWKKYSILTAVAKTLKEDTMAKOFLEBAALMKKHPNLVQ 446
QY 68 FYGVILEPPNYGIVTEVASLSGLYDINSNREEMDMHMTWATDVAKGHHYLMHEAPV 127
DB 447 LIGVCHREPPFYITIEPMHSHGULDFLRSGRETIDAVALLMYATQIAGSGYLSER--- 503
QY 128 KYIHRDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGT-PP-WMAPEVIGSLPVS 184
DB 504 NYIHRDLAARNCLVGENHVAVADFGSLRMTGDTYTAHAGAKFPKMTAPGALYNKPS 563

```


QY	185	ETCPTVAGVLLHMEVLTREV-PFKCLEGLQVAMLVENKNEPLTPSSCPRSPAEILHQCW	243
Db	564	TKSDVMAFGVLLMEATATGMSPPYPAID-LTVDYHKLDDGYMERBPCCPPEVYDLMRCQW	622
QY	244	EADAKRBPSPKQIISILESMGNDTSLPDKCNSFLHNKAEWCEIEATLERL-----	294
Db	623	QMDATDRPTFTFSIHHLAHEMFQESITBAVEKQINANNATSSASSAPSTSGVATGGATT	682
QY	295	-----KLLERDLSFKQEGLKERRERLLKMEQKLT-----EGSNTPLLLPLA	335
Db	683	TAAAGCASSSSATASLTLPDMVKKG-----LPGGALTPNAHNDPHQOQASTP-----	732
QY	336	ARMSEESYFEESKTEBSNAEMSCQATNSNEGHGMNLSLOAMLMGSGDILFSNMKAGAV	395
Db	733	--MSSTG-----STSTKLS--TFSSQCKGN-----	753
QY	396	MHSGMQINMQAKQ-----NSSKTTGKRRGKVNMALGFSDPLSEGD-----	437
Db	754	VQMRTTKKQCKQAAPPKRTSLSSSDSTYREEDPANANCFID-DLSTINGLARDINS	812
QY	438	-----DDDDDDGDEEDND	450
Db	813	LTRYDSETPDPAEDPTD	830
RESULT 14			
RYKL	DICD1	STANDARD;	PRT; 1584 AA.
ID	AC	P18160;	
DT	01-NOV-1990	(Rel. 16, Created)	
DT	01-OCT-1996	(Rel. 34, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) ('tyrosine-protein kinase 1').		
GN	PYKA OR SPLA OR DPYK1.		
OS	Dictyostelium discoideum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.		
OX	NCBI_TaxID=44689;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=JH10;		
RX	MEDLINE=97053827; PubMed=8898241;		
RA	Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;		
RT	"The Dictyostelium dual-specificity kinase splA is essential for spore differentiation.";		
RL	Development 122:3295-3305(1996).		
RN	[2]		
RP	SEQUENCE OF 1248-1584 FROM N.A.		
RX	MEDLINE=90287147; PubMed=1972546;		
RA	Tan J.L., Spudich J.A.;		
RT	"Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum.";		
RL	Mol. Cell. Biol. 10:3578-3583(1990).		
CC	-1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.		
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.		
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK DURING THE MOUND STAGE OF MORPHOGENESIS.		
CC	-1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.		
CC	-1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.		
CC	CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb.ch/announce/ or send an email to license@isb.ch).		
DR	EMBL; U32174; AAB41125.1; -		
DR	EMBL; M33785; AAA33202.1; -		
DR	PIR; T18276; T18276.		
DR	DictyDb; DD03010; pyKA.		

DR	InterPro; IPR0000719; Prot_kinase.
DR	InterPro; IPR001660; SAM.
DR	InterPro; IPR003877; SPRY_receptor.
DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00069; pkinase; 1.
DR	Pfam; PF00536; SAM; 1.
DR	Pfam; PF00622; SPRY; 3.
DR	PRINTS; PR00109; TYRKINASE.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	SMART; SM00454; SAM; 1.
DR	SMART; SM00449; SPRY; 3.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR	PROSITE; PS50105; SAM_DOMAIN; 1.
KW	Transferase; Tyrosine-Protein kinase; ATP-Binding; Phosphorylation.
FT	DOMAIN 908 972 SAM.
FT	DOMAIN 403 420 POLY-ASN.
FT	DOMAIN 428 435 POLY-THR.
FT	DOMAIN 449 480 POLY-ASN.
FT	DOMAIN 483 491 POLY-ASN.
FT	DOMAIN 494 508 POLY-ASN.
FT	DOMAIN 512 532 POLY-ASN.
FT	DOMAIN 596 600 POLY-ASN.
FT	DOMAIN 808 811 POLY-PHE.
FT	DOMAIN 1026 1029 POLY-SER.
FT	DOMAIN 1195 1210 POLY-ASN.
FT	DOMAIN 1215 1220 POLY-GLN.
FT	DOMAIN 1224 1233 POLY-GLN.
FT	DOMAIN 1266 1274 POLY-PRO.
FT	DOMAIN 1289 1561 PROTEIN_KINASE.
FT	NP_BIND 1295 1303 ATP (BY SIMILARITY).
FT	ACT_SITE 1316 1316 ATP (BY SIMILARITY).
FT	ACT_SITE 1417 1417 BY SIMILARITY.
FT	CONFLICT 1248 1249 D -> R (IN REF. 2).
FT	CONFLICT 1435 1435 V -> L (IN REF. 2).
Q	SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

	Query Match	16.9%	Score 401.5	DB 1	Length 1584
	Best Local Similarity	34.6%	Pred. No. 2.6e-18		
	Matches 102	Conservative 52	Mismatches 104	Indels 37	Gaps 8
Qy	10 QIKEDDLOFENNCGGSGFSGSYRAKWTISDQEKAVK-----KLTKIEKEAEILS 58				
Db	1283 EIDPNELEFGQTIGKGFPEVKRGY--RETDAVAKIITYRDQFKTKSLVAFQNEVGILS 1340				
Qy	59 VLSHRNIIOFYGVILE--PRNYGVTEYSALGSLYDVIINSRSEEMMDHI-MTWADVA 115				
Db	1341 KLRPNVAVQFAGCTAGGEBDHCHIVTEWGGSGSLRPLTHDFNLLLEONPHRLKLALDIA 1400				
Qy	116 KGMHYLMEAPVVKVIHRDLKSRNVI-----AADGVAKICDGFASRFHNHTT 162				
Db	1401 KGMNVLHGMPR-PIHRDLSRRNILLDHNIDPKNPVSSRODICKKISDFGLSKKEQA 1455				
Qy	163 HM-SLVGTFRPMAPAEVYQISLPVSETDITYSYGVVLMEMLTREVPFKGLQVAMLYVE 220				
Db	1460 SQMTQSGCIPYMAPEVFKGDSNSEKADVYSYGVVLEFELLTSDPEQDMKMKAKHALLAY 1519				
Qy	221 KNERLTIPSSCPRSFPAELLHQCEWADAKKRPSPFKQIISILESM-----SNDTSLP 270				
Db	1520 ESYRPIPIPLTTSSKMKELIYQCWMSNDSDRTPFQIILVHLKEMWDQGVSSFASYVP 1574				
RESULT 15					
ABL_FSVHY	ABL_FSVHY	STANDARD;	PRT;	439 AA.	
AC	PI0447;				
DT	01-MAR-1989 (Rel. 10, Created)				
DT	01-MAR-1989 (Rel. 10, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).				
SN	V-ABL.				
OS	Feline sarcoma virus (strain Hardy-Zuckerman 2).				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:11:58 ; Search time 40.8333 Seconds

(without alignments)
2875.443 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375
Sequence: 1 MSLSAFQIKFDDQFFE.....GDDDDDDGEEDNDNDNSE 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2375	100.0	455	4	Q9HCC4	Q9HCC4 homo sapien
2	2299.5	96.8	454	11	Q9ESL3	Q9ESL3 mus musculu
3	1751.5	73.7	800	4	Q9NYL2	Q9NYL2 homo sapien
4	1749.5	73.7	800	4	Q9HCC5	Q9HCC5 homo sapien
5	1748.5	73.6	800	4	Q9NYE9	Q9NYE9 homo sapien
6	1748.5	73.6	800	4	Q9HDD2	Q9HDD2 homo sapien
7	1743.5	73.4	802	11	Q9ESL4	Q9ESL4 mus musculu
8	1481	62.4	289	11	Q8BR73	Q8BR73 mus musculu
9	1402.5	59.1	371	13	Q9QZV8	Q9QZV8 brachydanto
10	600	25.3	1066	4	Q9H2N5	Q9H2N5 homo sapien
11	600	25.3	608	11	Q8B1G8	Q8B1G8 mus musculu
12	598.5	25.2	1001	11	Q8YD66	Q8YD66 mus musculu
13	597.5	25.2	564	4	Q9H1Y7	Q9H1Y7 homo sapien
14	590.5	24.9	1036	4	Q8WMN1	Q8WMN1 homo sapien
15	588.5	24.8	570	4	Q8WMN2	Q8WMN2 homo sapien
16	581.5	24.5	847	4	Q16584	Q16584 homo sapien

17	580.5	24.4	850	11	Q9J15	Q9J15 mus musculu
18	569.5	24.0	968	11	Q8CDL6	Q8CDL6 mus musculu
19	568.5	23.9	886	4	Q43283	Q43283 homo sapien
20	565.5	23.8	859	4	Q8WY25	Q8WY25 mus sapien
21	561.5	23.6	888	11	Q8CBX3	Q8CBX3 mus musculu
22	560.5	23.6	406	10	Q23719	Q23719 arabidopsis
23	560.5	23.6	880	10	Q859K4	Q859K4 arabidopsis
24	544.5	22.9	1148	5	Q95VP6	Q95VP6 drosophila
25	544.5	22.9	1161	5	Q95UN8	Q95UN8 drosophila
26	535.5	22.5	1161	5	Q8WRK7	Q8WRK7 drosophila
27	529	22.3	886	10	Q9LY18	Q9LY18 arabidopsis
28	527.5	22.2	977	5	Q9VW24	Q9VW24 drosophila
29	510.5	21.5	1030	10	Q8L625	Q8L625 arabidopsis
30	510.5	21.5	1030	10	Q9C9U5	Q9C9U5 arabidopsis
31	507.5	21.4	109	11	Q8BUD3	Q8BUD3 mus musculu
32	505	21.3	855	5	Q01700	Q01700 caenorhabd
33	495	20.8	982	10	Q65833	Q65833 lycopersico
34	494	20.8	903	10	Q9FPR5	Q9FPR5 oryza sativ
35	483.5	20.4	847	10	Q93XL9	Q93XL9 rosa hybrid
36	483	20.3	491	4	Q9NZ70	Q9NZ70 homo sapien
37	480	20.2	829	10	Q24027	Q24027 lycopersico
38	480	20.2	992	10	Q8LPH3	Q8LPH3 arabidopsis
39	478.5	20.1	933	10	Q9FPR3	Q9FPR3 arabidopsis
40	475.5	20.0	412	10	Q9M085	Q9M085 arabidopsis
41	475	20.0	957	10	Q9FPR4	Q9FPR4 hordeum vul
42	472.5	19.9	616	13	Q73613	Q73613 xenopus lae
43	471.5	19.9	606	11	Q923A8	Q923A8 mus musculu
44	471	19.8	777	10	Q9C833	Q9C833 arabidopsis
45	471	19.8	806	10	Q9ZSD8	Q9ZSD8 lycopersico

ALIGNMENTS

RESULT 1

ID Q9HCC4 PRELIMINARY; PRT; 455 AA.
AC Q9HCC4;
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE MKTK-beta (Similar to sterile-alpha motif and leucine zipper
DE containing kinase Azk) (Mixed lineage kinase) (Mixed lineage kinase-
DE related kinase MKK-beta).
GN MKTK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MKTK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Acton S.;
RT "MK-mixed lineage kinase.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP MEDLINE=21950776; PubMed=11836244;
RA Gross E.A., Callow W.G., Waldbaum U., Thomas S., Ruggieri R.;
RT "WRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in
RT gamma-Radiation-induced Cell Cycle Arrest.";
RL J. Biol. Chem. 277:13873-13882(2002).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AB049734; BAB16445.1; -
 DR EMBL: BC001401; AAH01401.1; -
 DR EMBL: AF325454; AAK1615.1; -
 DR EMBL: AF480462; AAL85892.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR InterPro: IPR001245; Tyr kinase.
 DR Pfam: PF00669; Kinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot kinase; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 DR ATP-binding: Kinase; Serine/threonine-protein kinase; Transferase.
 KM
 SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 2375; DB 4; Length 455;
 Best Local Similarity 100.0%; Pred. No. 3.7e-180;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGASFYQIKFDDLOFENCGGSGFSGYRAKWIISQKEVAVKKLTKEAEILSVL 60
 DB 1 MSLSGASFYQIKFDDLOFENCGGSGFSGYRAKWIISQKEVAVKKLTKEAEILSVL 60
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNSEEMDMHIMTATDVAKGMY 120
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNSEEMDMHIMTATDVAKGMY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASRFNHTHTMSLVGTFPMMAPEVIO 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASRFNHTHTMSLVGTFPMMAPEVIO 180
 QY 181 LPSVETCDTYSYGVVLEMTREVPFKGLEGLQVAMLVYKNERLTTPSCPSFAELH 240
 DB 181 LPSVETCDTYSYGVVLEMTREVPFKGLEGLQVAMLVYKNERLTTPSCPSFAELH 240
 QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATELRLKLEND 300
 DB 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATELRLKLEND 300
 QY 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTESNSAEMSCOI 360
 DB 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTESNSAEMSCOI 360
 QY 361 TATSNEBGHGMNPSIQAMMLMGFDIFSNKAGAVHSGMOINMOAKONSSKTTCKRCK 420
 DB 361 TATSNEBGHGMNPSIQAMMLMGFDIFSNKAGAVHSGMOINMOAKONSSKTTCKRCK 420
 QY 421 KYNMAGFSDPDLSEGGDDDDGEEEDNDMDNSE 455
 DB 421 KYNMAGFSDPDLSEGGDDDDGEEEDNDMDNSE 455

RESULT 2

Q9ESL3 PRELIMINARY; PRT; 454 AA.
 AC Q9ESL3;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE MLTK-beta.
 GN ZAK OR MLTK.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21264927; PubMed=11042189;
 RA Gotoh I., Aachi M., Nishida E.,
 RT "Identification and Characterization of a Novel MAP Kinase Kinase
 RT Kinase, MLTK.";
 RL J. Biol. Chem. 276:4276-4286(2001).

CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AB049732; BAB16443.1; -
 DR HSSP: P12931; 1FMK.
 DR GMD; MG1:1931274; Zak.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR InterPro: IPR001245; Tyr kinase.
 DR Pfam: PF00669; Kinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot kinase; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 DR ATP-binding: Kinase; Serine/threonine-protein kinase; Transferase.
 KM
 SQ SEQUENCE 454 AA; 51366 MW; 35C2FCDD729D395 CRC64;

Query Match 96.8%; Score 2299.5; DB 11; Length 454;
 Best Local Similarity 96.5%; Pred. No. 3.6e-174;
 Matches 439; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

QY 1 MSLSGASFYQIKFDDLOFENCGGSGFSGYRAKWIISQKEVAVKKLTKEAEILSVL 60
 DB 1 MSLSGASFYQIKFDDLOFENCGGSGFSGYRAKWIISQKEVAVKKLTKEAEILSVL 60
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNSEEMDMHIMTATDVAKGMY 120
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNSEEMDMHIMTATDVAKGMY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASRFNHTHTMSLVGTFPMMAPEVIO 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASRFNHTHTMSLVGTFPMMAPEVIO 180
 QY 181 LPSVETCDTYSYGVVLEMTREVPFKGLEGLQVAMLVYKNERLTTPSCPSFAELH 240
 DB 181 LPSVETCDTYSYGVVLEMTREVPFKGLEGLQVAMLVYKNERLTTPSCPSFAELH 240
 QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATELRLKLEND 300
 DB 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATELRLKLEND 300
 QY 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTESNSAEMSCOI 360
 DB 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTESNSAEMSCOI 360
 QY 361 TATSNEBGHGMNPSIQAMMLMGFDIFSNKAGAVHSGMOINMOAKONSSKTTCKRCK 420
 DB 361 TATSNEBGHGMNPSIQAMMLMGFDIFSNKAGAVHSGMOINMOAKONSSKTTCKRCK 420
 QY 421 KYNMAGFSDPDLSEGGDDDDGEEEDNDMDNSE 455
 DB 421 KYNMAGFSDPDLSEGGDDDDGEEEDNDMDNSE 455

RESULT 3

Q9NYL2 PRELIMINARY; PRT; 800 AA.
 AC Q9NYL2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Mixed lineage kinase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20384179; PubMed=10924358;
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,
 RA Chou C.K., Yang J.J.;
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein
 RT containing a leucine-zipper and a sterile-alpha motif.";
 RL Biochem. Biophys. Res. Commun. 274:811-816(2000).
 CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AF238255; AAF63490.1; -.
 DR HSSP; P12931; IFMK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001660; SAM_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50105; SAM DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 73.7%; Score 1751.5; DB 4; Length 800;
 Best Local Similarity 76.1%; Pred. No. 2.5e-130;
 Matches 360; Conservative 24; Mismatches 56; Indels 33; Gaps 9;

QY 1 MSLSGASFYQIKFDDIQFENCGGSGFSGSVYRAKWIISQKEVAVKLLIKEAEILSVL 60
 |||||
 Db 1 MSLSGASFYQIKFDDIQFENCGGSGFSGSVYRAKWIISQKEVAVKLLIKEAEILSVL 60
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSRSEEMDMHIMTWATDVAKGMHY 120
 |||||
 Db 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSRSEEMDMHIMTWATDVAKGMHY 120
 QY 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 |||||
 Db 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 QY 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 |||||
 Db 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 QY 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 |||||
 Db 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 QY 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 |||||
 Db 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 QY 241 OCMEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAERCEIATELRLKLERD 300
 |||||
 Db 241 OCMEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAERCEIATELRLKLERD 300
 QY 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 |||||
 Db 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 QY 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 |||||
 Db 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 QY 356 MSCQIATNSGEGHGNPISLOAMLMFGFDIFSMN--KAGAVNH--SGMQ-----INM- 404
 |||||
 Db 356 MSCQIATNSGEGHGNPISLOAMLMFGFDIFSMN--KAGAVNH--SGMQ-----INM- 404
 QY 361 MSVYASLFFKEN-----NITGKRLLILLEEDLDKMGIVSGHIIHFSAIETKLTHTDYNLF 415
 |||||
 Db 361 MSVYASLFFKEN-----NITGKRLLILLEEDLDKMGIVSGHIIHFSAIETKLTHTDYNLF 415
 QY 405 ---QAQONSSKTSKRGRKKNVMAALGFSDPDLSEGDGDDDD-----DGEER 446
 |||||
 Db 405 ---QAQONSSKTSKRGRKKNVMAALGFSDPDLSEGDGDDDD-----DGEER 446
 QY 416 HPPPLIKDGGGEENEKEKIVNLELVFG-FHLKPGTGPODCKWRYMEMDGD 467
 |||||
 Db 416 HPPPLIKDGGGEENEKEKIVNLELVFG-FHLKPGTGPODCKWRYMEMDGD 467

RESULT 4

Q9HCC5 PRELIMINARY; PRT; 800 AA.
 ID Q9HCC5
 AC Q9HCC5;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE MLTK-alpha.
 GN MLTK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21264927; PubMed=11042189;
 RA Gotoh I.; Adachi M.; Nishida E.;
 RT "Identification and Characterization of a Novel MAP Kinase Kinase, MLTK.";

RL J. Biol. Chem. 276:4276-4286 (2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AB049733; BAB16444.1; -.
 DR HSSP; P12931; IFMK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001660; SAM_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50105; SAM DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FDD CRC64;

Query Match 73.7%; Score 1749.5; DB 4; Length 800;
 Best Local Similarity 76.1%; Pred. No. 3.6e-130;
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSLSGASFYQIKFDDIQFENCGGSGFSGSVYRAKWIISQKEVAVKLLIKEAEILSVL 60
 |||||
 Db 1 MSLSGASFYQIKFDDIQFENCGGSGFSGSVYRAKWIISQKEVAVKLLIKEAEILSVL 60
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSRSEEMDMHIMTWATDVAKGMHY 120
 |||||
 Db 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSRSEEMDMHIMTWATDVAKGMHY 120
 QY 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 |||||
 Db 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 QY 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 |||||
 Db 121 LHMEAVKVIHRDLKSRNVIAADGVLCIDFGASFHNHTTMSLVGTFPMNAPEVIO 180
 QY 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 |||||
 Db 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 QY 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 |||||
 Db 181 LPVSECTDYTSYGVVLMEMLTREVPPKGLGLOVAVLVKNERLITPSSCPSPFALLH 240
 QY 241 OCMEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAERCEIATELRLKLERD 300
 |||||
 Db 241 OCMEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAERCEIATELRLKLERD 300
 QY 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 |||||
 Db 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 QY 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 |||||
 Db 301 LSFKEQLERERRRLKMWEOQLTEQNTPL--LPLAAMSESYF---ESKTEENSAR 355
 QY 356 MSCQIATNSGEGHGNPISLOAMLMFGFDIFSMN--KAGAVNH--SGMQ-----INM- 404
 |||||
 Db 356 MSCQIATNSGEGHGNPISLOAMLMFGFDIFSMN--KAGAVNH--SGMQ-----INM- 404
 QY 361 MSVYASLFFKEN-----NITGKRLLILLEEDLDKMGIVSGHIIHFSAIETKLTHTDYNLF 415
 |||||
 Db 361 MSVYASLFFKEN-----NITGKRLLILLEEDLDKMGIVSGHIIHFSAIETKLTHTDYNLF 415
 QY 405 ---QAQONSSKTSKRGRKKNVMAALGFSDPDLSEGDGDDDD-----DGEER 446
 |||||
 Db 405 ---QAQONSSKTSKRGRKKNVMAALGFSDPDLSEGDGDDDD-----DGEER 446
 QY 416 HPPPLIKDGGGEENEKEKIVNLELVFG-FHLKPGTGPODCKWRYMEMDGD 467
 |||||
 Db 416 HPPPLIKDGGGEENEKEKIVNLELVFG-FHLKPGTGPODCKWRYMEMDGD 467

RESULT 5

Q9NYE9 PRELIMINARY; PRT; 800 AA.
 ID Q9NYE9
 AC Q9NYE9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed lineage kinase-related kinase MKR-alpha).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX McNeel J.J.; Frima N.; Diamond T.E.; Power S.K.; Guesdon F.;
 RT "Cloning and Characterisation of AZK, a mixed lineage kinase

RT containing a sterile-alpha motif";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE:21950776; PubMed:11836244;
 RX Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;
 RT "MRK, a Mixed-lineage Kinase-related Molecule That Plays a Role in
 RT gamma-Radiation-induced Cell Cycle Arrest";
 RL J. Biol. Chem. 277:13873-13882(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF251441; AAF65822.1; -
 DR EMBL: AF480461; AAL85891.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF00536; SAM; 1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00454; SAM; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE_ST; 1.
 DR PROSITE: PS50105; SAM DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 800 AA; 91181 MW; B289D836C5E2E295 CRC64;
 SQ
 Query Match 73.6%; Score 1748.5; DB 4; Length 800;
 Best Local Similarity 76.1%; Pred. No. 4.3e-130;
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;
 QY 1 MSSLGASFVQIKFDDLOFFENCSSGSGSVYRAKWI SODKEVAVKLLKIEKEAELISVL 60
 DB 1 MSSLGASFVQIKFDDLOFFENCSSGSGSVYRAKWI SODKEVAVKLLKIEKEAELISVL 60
 QY 61 SHRNIIQFYGVILEPPNNGIYEVASLSGLYDYINSNSEMDMHIWMTATDVAKGMY 120
 DB 61 SHRNIIQFYGVILEPPNNGIYEVASLSGLYDYINSNSEMDMHIWMTATDVAKGMY 120
 QY 121 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFNHTTHMSLVGTFPMMAPEVIO 180
 DB 121 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFNHTTHMSLVGTFPMMAPEVIO 180
 QY 181 LPSVETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLTISSCPSPFAELH 240
 DB 181 LPSVETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLTISSCPSPFAELH 240
 QY 241 OCWEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLKLERD 300
 DB 241 OCWEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLKLERD 300
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPL--LPLAARMSSESYF---ESKTEESNSAE 355
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPL--LPLAARMSSESYF---ESKTEESNSAE 355
 QY 360 LSFKEQELKERERRRLKMEQKLTQOSNTPLPSFEIGAWTEDDVYCWVQOLVRKGDSSAE 360
 DB 360 LSFKEQELKERERRRLKMEQKLTQOSNTPLPSFEIGAWTEDDVYCWVQOLVRKGDSSAE 360
 QY 361 MSCQITATNSNGEHCNPNLSQAMMLMGFGDIFSN--KAGAVMH--SGNQ-----INN- 404
 DB 361 MSCQITATNSNGEHCNPNLSQAMMLMGFGDIFSN--KAGAVMH--SGNQ-----INN- 404
 QY 405 ---QAKNSSKTTSKRGRKKNMVALGFSDFDLSGGDDDD-----DGGEE 446
 DB 405 ---QAKNSSKTTSKRGRKKNMVALGFSDFDLSGGDDDD-----DGGEE 446
 QY 416 HPPPLIKDSGGEPENBEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGD 467
 DB 416 HPPPLIKDSGGEPENBEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGD 467

RESULT 6

Q9HDD2 PRELIMINARY; PRT; 800 AA.
 AC Q9HDD2;
 DT 01-MAR-2001 (Tremblrel). 16, Created)
 DT 01-MAR-2001 (Tremblrel). 16, last sequence update)
 DT 01-MAR-2003 (Tremblrel). 23, last annotation update)
 DE Plausible mixed-lineage kinase protein.

GN MLKXAK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lymphoid organ;
 RA Abe Y., Ueda N.;
 RT "Placible Mixed-lineage kinase derived from LAK cell.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AB030034; BAB12040.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF00536; SAM; 1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00454; SAM; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE_ST; 1.
 DR PROSITE: PS50105; SAM DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 800 AA; 91155 MW; B2814509EC54B07A CRC64;
 SQ
 Query Match 73.6%; Score 1748.5; DB 4; Length 800;
 Best Local Similarity 76.1%; Pred. No. 4.3e-130;
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;
 QY 1 MSSLGASFVQIKFDDLOFFENCSSGSGSVYRAKWI SODKEVAVKLLKIEKEAELISVL 60
 DB 1 MSSLGASFVQIKFDDLOFFENCSSGSGSVYRAKWI SODKEVAVKLLKIEKEAELISVL 60
 QY 61 SHRNIIQFYGVILEPPNNGIYEVASLSGLYDYINSNSEMDMHIWMTATDVAKGMY 120
 DB 61 SHRNIIQFYGVILEPPNNGIYEVASLSGLYDYINSNSEMDMHIWMTATDVAKGMY 120
 QY 121 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFNHTTHMSLVGTFPMMAPEVIO 180
 DB 121 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFNHTTHMSLVGTFPMMAPEVIO 180
 QY 181 LPSVETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLTISSCPSPFAELH 240
 DB 181 LPSVETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLTISSCPSPFAELH 240
 QY 241 OCWEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLKLERD 300
 DB 241 OCWEADAKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLKLERD 300
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPL--LPLAARMSSESYF---ESKTEESNSAE 355
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPL--LPLAARMSSESYF---ESKTEESNSAE 355
 QY 360 LSFKEQELKERERRRLKMEQKLTQOSNTPLPSFEIGAWTEDDVYCWVQOLVRKGDSSAE 360
 DB 360 LSFKEQELKERERRRLKMEQKLTQOSNTPLPSFEIGAWTEDDVYCWVQOLVRKGDSSAE 360
 QY 361 MSCQITATNSNGEHCNPNLSQAMMLMGFGDIFSN--KAGAVMH--SGNQ-----INN- 404
 DB 361 MSCQITATNSNGEHCNPNLSQAMMLMGFGDIFSN--KAGAVMH--SGNQ-----INN- 404
 QY 405 ---QAKNSSKTTSKRGRKKNMVALGFSDFDLSGGDDDD-----DGGEE 446
 DB 405 ---QAKNSSKTTSKRGRKKNMVALGFSDFDLSGGDDDD-----DGGEE 446
 QY 416 HPPPLIKDSGGEPENBEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGD 467
 DB 416 HPPPLIKDSGGEPENBEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGD 467

RESULT 7

Q9ESL4 PRELIMINARY; PRT; 802 AA.
 AC Q9ESL4;
 DT 01-MAR-2001 (Tremblrel). 16, Created)
 DT 01-MAR-2001 (Tremblrel). 16, last sequence update)
 DT 01-MAR-2003 (Tremblrel). 23, last annotation update)

DE MLTK alpha (Sterile-alpha motif and leucine zipper containing kinase
 DE AZK).
 GN ZAK OR MLTK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxId=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21264927; PubMed=11042189;
 RA Gotoh I., Adachi M., Nishida E.;
 RT "Identification and Characterization of a Novel MAP Kinase Kinase
 RT Kinase, MLTK.";
 RL J. Biol. Chem. 276:4276-4286(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AB049731; BAB16442.1; -;
 DR EMBL; BC023718; AAH23718.1; -;
 DR HSSP; P12931; IFMK.
 DR MGD; MGI:1931274; Zak.
 DR InterPro; IPR000194; ATPase a/bcentre.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00536; SAM; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00454; SAM; 1_kinase; 1.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA_1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS0105; SAM_DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 802 AA; 91719 MW; D431DF8312A43C CRC64;

Query Match 73.4%; Score 1743.5; DB 11; Length 802;
 Best Local Similarity 74.3%; Pred. No. 1.1e-129;
 Matches 356; Conservative 20; Mismatches 58; Indels 45; Gaps 7;

QY 1 MSLSGASFOVQIKRDDIQFENCGGSGFSGVYRAKWTISQKEVAVKKLKTEKEAEITLSVL 60
 DB 1 MSLSGASFOVQIKRDDIQFENCGGSGFSGVYRAKWTISQKEVAVKKLKTEKEAEITLSVL 60
 QY SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 DB SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 QY 61 SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 DB 61 SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 DB 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 DB 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 QY 181 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 DB 181 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 QY 241 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 DB 241 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 QY 241 OCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSTFLHNKAERCEIATEILKLEKLERD 300
 DB 241 OCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSTFLHNKAERCEIATEILKLEKLERD 300
 QY 301 LSFKEQLKERERRLKMEQKLTQOSTPL--LPLAARMSSESYF-----ESKTE 349
 DB 301 LSFKEQLKERERRLKMEQKLTQOSTPL--LPLAARMSSESYF-----ESKTE 349
 QY 301 LSFKEQLKERERRLKMEQKLTQOSTPL--LPLAARMSSESYF-----ESKTE 349
 DB 301 LSFKEQLKERERRLKMEQKLTQOSTPL--LPLAARMSSESYF-----ESKTE 349
 QY 350 ESNASAMSCITATNSGEGHGNPQLQAMMLMGFGDIFSN--KAGAVHSGMQINQAK 407
 DB 361 MSGVYALIFKENNTIG-----KRLLLLEEEDLKDMGIVSKGHIHFKAIEKLTH 409
 QY 408 -----ONSKTTSKRGRKKNMVLGFSDFLISGDDDD-----DDGEE 446
 DB 408 -----ONSKTTSKRGRKKNMVLGFSDFLISGDDDD-----DDGEE 446

DB 410 DYINLFHPPELIKDSGEBENEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGE 467
 RESULT 8
 ID Q9BR73 PRELIMINARY; PRT; 289 AA.
 AC Q9BR73;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Sterile-alpha motif and leucine zipper containing kinase AZK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK045444; BAC32371.1; -;
 SQ SEQUENCE 289 AA; 32872 MW; 550E5651066A0463 CRC64;

Query Match 62.4%; Score 1481; DB 11; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.8e-109;
 Matches 280; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLSGASFOVQIKRDDIQFENCGGSGFSGVYRAKWTISQKEVAVKKLKTEKEAEITLSVL 60
 DB 1 MSLSGASFOVQIKRDDIQFENCGGSGFSGVYRAKWTISQKEVAVKKLKTEKEAEITLSVL 60
 QY SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 DB SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 QY 61 SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 DB 61 SHRNIIQFQGVILEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMY 120
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 DB 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 DB 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFNHTTHMSLVGTFPMWAPVIOG 180
 QY 181 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 DB 181 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 QY 241 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 DB 241 LPVSETCDYISGVVLMEMLTREVPPKGLGEGVAMLVKXKRLTIPSSCPSPFALLH 240
 QY 241 OCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSTFLHNKAER 284
 DB 241 OCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSTFLHNKAER 284

RESULT 9
 ID Q90ZV8 PRELIMINARY; PRT; 371 AA.
 AC Q90ZV8;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Brachydanio rerio (Zebrafish) (Danio rerio).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 RN NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
 RT "A novel protein kinase, zNPK, from the zebrafish.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF265343; AAK52416.1; -;
 DR InterPro; IPR000719; Prot_kinase.


```

Db      245 GNNVLDHEAIVPIIHRDLKSSNILLQKVENGDLSNKILKTIDFGIAREWHRTTKMSAG 304
Qy      169 TFFWMAPEVYQSIPVSETCCTVYGVLMEMLTREVPFKLEGLQVAMLVENKEELTIP 228
Db      305 TYAMMAPEVIRASMFSGSVWSYGVLMELLGVEVFRIDGLAAVAVANMKALPLP 364
Qy      229 SSCPRFAELHOCWEADAKKRPSPFOIISILSMUSDLSLPDKNSFLINKAEMCEIE 288
Db      365 STPEEPFAKMECQMPDPHSRPSFTSIDQLTTIESGFPFMPKPSFHLQDQMKHEIQ 424
Qy      289 ATERLKLEKRLDSFKEQEL-----KERERLKMWEQKLTESQ----- 326
Db      425 EMEDQIRAKEKEKRLTWEELTRAALOQKQEEILRRREQELAREIDILEREINIIHQ 484
Qy      327 -----NTPLLLPLAARSESYFESKTEESNSAEMSCQITATNGEGHGM 371
Db      485 QGKRPVKKRKGKFRKSRRLKDKGNRLSLPSDFQHKFTVQASPTMKRKLINRSPPA 544
Qy      372 NPS-----IQMMIMGFQDIFSNKAGAVMHSQMIMQAKONSSKTTSKRGKVMALG 427
Db      545 SPTIIRLRALQLTPESSKTWGRSSVY-----PKEGEERKRAPKKKGRTWG 593
Qy      428 ---FSDPDLSEGD 438
Db      594 PGTIGQKELTSGDE 607

```

RESULT 12

```

Q8VDG6 PRELIMINARY; PRT; 1001 AA.
ID Q8VDG6
AC Q8VDG6;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase 9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; BC021891; AH421891.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase.1
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1001 AA; 109983 MW; E10042C868B953C CRC64;

```

Query Match 25.2%; Score 598.5; DB 11; Length 1001;

Best Local Similarity 33.7%; Pred. No. 1,le-38;

Matches 161; Conservative 85; Mismatches 161; Indels 71; Gaps 16;

```

Qy      9 VQKFDLDQFENCGGSGSVYRAKTIQDKVAVYKLLK-----IEKRAEI 56
Db      103 VHVDFELRLKEKILIGAGFGQVYRATV--OGQEVAVVAARDPEQDAAAAESVRRARL 160
Qy      57 LSVLSHRNIIQFVGLIEPPNYGIVTEYASLGIYDI-----NSNRSEMDMDH 106

```

```

Db      161 FAWLRHPIIQLRGVCLRPQHLCLVLEFARGALNRALAAASDPRAPGRARRIPQV 220
Qy      107 IMTWADVAKGMYHYLMERPVKVIHRDLKSRNVV-----AAGVYLKICDFGASRFH 158
Db      221 LVNMAVOAIRGMVLYLHEEAVVPLIHRDLKSSNILLLEKIEHDDICNKTLLKIDPFGIAREW 280
Qy      159 NHTTHSLVGTFFWMAPEVYQSIPVSETCCTVYGVLMEMLTREVPFKLEGLQVAMLV 218
Db      281 HRTTRMSAAGTYAMMAPEVIRASLFSKGSIDINSYGVLMELLGVEVFRIDGLAAVAVGY 340
Qy      219 VEKNERLTIPSSCPSPFAELHOCWEADAKKRPSPFOIISILSMUSDLSLPDKNSP 276
Db      341 AVNKLFLPSTCPPEPFAKMECQDPHRSFALLIQQLTAIEAVLTNNPQE--SF 398
Qy      277 LNKAWRCIEATLERLKLKERDLSFKEQEL-----KERERLKMWEQKLTESQNTP 329
Db      399 HSNQEDWKLEIQFSS-LRTKEKELRSREELSRALLOQKSOBELLRREQOLAREIDIV 457
Qy      330 L--LPLAARSESYFESKTEESNSAEMSCQITATNGEGHGM-PS-----LQA 377
Db      458 LERELVLLFQLSQEPHYKKRGRFRGRRLK-----DGRISLPSDFQHKITVQA 510
Qy      378 MMLMGFDIFSNKAGAVMHSQ---MQIMQAKONSSKTTSKRGKVMALGFSPD 432
Db      511 SPTL---DKRRSDSGLCPGSPPLMLPRLRALQLTSDENNKTGR--NVVFRQEDPE 563

```

RESULT 13

```

Q9HIY7 PRELIMINARY; PRT; 564 AA.
ID Q9HIY7
AC Q9HIY7;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE D7862P8.3 (Similar to MAP3K10 (Mitogen-activated protein kinase kinase 10)) (Fragment).
DE D7862P8.3.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hall R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AL133380; CAC17571.1; -.
DR HSSP; P06241; ISHF.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KM ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase; Transferase.
SQ NON_TER 564
SQ SEQUENCE 564 AA; 62400 MW; D4F91EL4842E7EEB CRC64;

```

Query Match 25.2%; Score 597.5; DB 4; Length 564;

Best Local Similarity 37.4%; Pred. No. 6,2e-39;

Matches 135; Conservative 66; Mismatches 115; Indels 45; Gaps 7;

```

QY 6 ASFVQIKFDLDLPFFENGCGGSGFSVYRAKTIQDKVAVKLLK-----IEKE 53
DB 114 SSSPVHVAFERLEIKELIGAGFGQVYRATW--QGQEVAVKARQDEQDAAAAAESVRE 171
QY 54 AELLVLSHRNIIIOFGVILEPPNYGIVTEYASLSGYDI-----NSNRSEE 101
DB 172 ARLFAMLRPNIIIELRGVLCIQPHLCVLEFARGALNRLAANAAPDRAPRRARR 231
QY 102 MDMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGVLKICDFG 153
DB 232 IPRHVLVNVAAVOIARGMVYLHEEAFVPIHRDLKSSNILLLEIKEDHDI CNKTLKITDFG 291
QY 154 ASRFNHTHTMSLVGTFPPMAPEVYQSLPVSECTDITYSGVVMEMLTREVPKGLGLO 213
DB 292 LAREWHRTTKMSTAGTYAMMAPEVYKSLFSKSDIWSGVLLMELTGEVPRGIDGLA 351
QY 214 VAMLVVEKNERLTIPSSCPRSFAELLHQCEADAKKRPSFKQIISLESMSND--TSLPD 271
DB 352 VAYGVAVNKLTLPISCTCPPEPAKLMKECWQDPHIRPSFALLLEQTLAEGAVMTEMPO 411
QY 272 KCONSPLHNKAEMRCIEATLERLKLKLERDISFKQEQL-----KERERLKMGEQKLTE 324
DB 412 E--SFHSMQDDWKLEIQWFDELRTKEKELRSREBELTRAALQOKSQEBLLKRREQOLAE 469
QY 325 Q 325
DB 470 R 470

```

RESULT 14

```

Q8WMN1 ID Q8WMN1 PRELIMINARY; PRT; 1036 AA.

```

```

AC 08WMN1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mixed lineage kinase 4beta.
GN MLK4BETA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kvaasna S., Protodopov A., Rynditch A., Zabarovsky E., Kashuba V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AJ311797; CAC84640.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001452; SH3_
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3_1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 1036 AA; 113809 MW; B9C2ACF397C2CFC37 CRC64;

```

Query Match 24.9%; Score 590.5; DB 4; Length 1036;
 Best Local Similarity 33.1%; Pred. No. 5,1e-38;
 Matches 149; Conservative 75; Mismatches 147; Indels 79; Gaps 9;

```

QY 6 ASFVQIKFDLDLPFFENGCGGSGFSVYRAKTIQDKVAVKLLK-----IEKE 53
DB 114 SSSPVHVAFERLEIKELIGAGFGQVYRATW--QGQEVAVKARQDEQDAAAAAESVRE 171
QY 54 AELLVLSHRNIIIOFGVILEPPNYGIVTEYASLSGYDI-----NSNRSEE 101
DB 172 ARLFAMLRPNIIIELRGVLCIQPHLCVLEFARGALNRLAANAAPDRAPRRARR 231
QY 102 MDMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGVLKICDFG 153
DB 232 IPRHVLVNVAAVOIARGMVYLHEEAFVPIHRDLKSSNILLLEIKEDHDI CNKTLKITDFG 291
QY 154 ASRFNHTHTMSLVGTFPPMAPEVYQSLPVSECTDITYSGVVMEMLTREVPKGLGLO 213
DB 292 LAREWHRTTKMSTAGTYAMMAPEVYKSLFSKSDIWSGVLLMELTGEVPRGIDGLA 351
QY 214 VAMLVVEKNERLTIPSSCPRSFAELLHQCEADAKKRPSFKQIISLESMSND--TSLPD 271
DB 352 VAYGVAVNKLTLPISCTCPPEPAKLMKECWQDPHIRPSFALLLEQTLAEGAVMTEMPO 411
QY 272 KCONSPLHNKAEMRCIEATLERLKLKLERDISFKQEQL-----KERERLKMGEQKLTE 324
DB 412 E--SFHSMQDDWKLEIQWFDELRTKEKELRSREBELTRAALQOKSQEBLLKRREQOLAE 469
QY 325 Q-----SNTPLLPLAARMESEYFESKTEESNA 354
DB 470 REIDVLRELNIILFQNLQEKPKVKRKGKFRSRILKLDGHRISLPDFQKITVOASP 529
QY 355 EM-----SCQITATSNQEGHGNPSLOAMML 380
DB 530 NDKRRSLNSSSSSPSSPTMPLRLATOL 559

```

RESULT 15

```

Q8WMN2 ID Q8WMN2 PRELIMINARY; PRT; 570 AA.

```

```

AC 08WMN2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mixed lineage kinase 4alpha.
GN MLK4ALPHA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kvaasna S., Protodopov A., Rynditch A., Zabarovsky E., Kashuba V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AJ311797; CAC84639.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001452; SH3_
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3_1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW Transferase.

```

SEQ SEQUENCE 570 AA; 62922 MW; 34CE937FCBCC5F93 CRC64;
Query Match 24.8%; Score 588.5; DB 4; Length 570;
Best Local Similarity 37.1%; Pred. No. 3.3e-38;
Matches 134; Conservative 66; Mismatches 116; Indels 45; Gaps 7;
QY 6 ASFVQIKFDDLOPFENCGGSGFSGSVYRAKMWISQDKEAVKLLK-----IENE 53
DB 114 SSPVAVAFERLEKEIGAGGQYVRAW--QGCEVAVKARQDPEQDAAAAAESVRE 171
QY 54 AELSVLSHRNIIQFYGVILEPPNYGIVTEVASLGSLYDI-----NSNRSEE 101
DB 172 ARLFALRLRHPNIIELRGVCLIQPHLCLVLEFARGALNRLAANAAPPRAAPRRARR 231
QY 102 MEMDHIMTWATDVAKGMHYLHMEAPYKVVIHRDLKSRNVI-----AADGLKICDFG 153
DB 232 IPPHVAVNNAVOJARGMVLYLHEEAFVPIILHRDLKSSNILLLEIEHDDICNKTLLKIDFG 291
QY 154 ASRFHNHTTHMSLVGTFPMWAPPEVIOSLPVSETCDTYSYGVVLMEMLTREVPFKGLEGO 213
DB 292 LAREMHTRTKMGAGTYAMMAPEVTKSSLPFSKGSIDIMSCGVILMELLTGEVPRGIDGLA 351
QY 214 VAMLVVEKNERLTISSCPRSPFAELLHQCEADAKRRPSFKOIIISLESMSND--TSLPD 271
DB 352 VAYGVAVNKLTLPISTCEPPEFAKMKECWQODPHIRPSFALILEQLTAIEGAVMTEMPO 411
QY 272 KNSFLHNKAEMRCIEATLERLKLERDLSFKEQL-----KERERLKMMEOKLTE 324
DB 412 E--SFHSMQDDWKLEIQOMFDELRTKEKELRSREBELTRAALQOKSQEELKRRREQOLAE 469
QY 325 Q 325
DB 470 R 470

Search completed: December 5, 2003, 09:20:15
Job time : 44.8333 secs

THIS PAGE BLANK (USP 100)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 08:30:40 ; Search time 49.2593 Seconds
(without alignments)
1466.132 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375

Sequence: 1 MSLGASFQIKFDLQFE.....GDDDDDDGDEEDNDNDSE 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2375	100.0	455	21	AA18657 A human regulator
2	2375	100.0	455	21	AA18657 Human survival reg
3	2375	100.0	455	21	AA18657 A human cardiovascular
4	2366	99.6	473	22	AA25322 Human protein sequ
5	1748.5	73.6	800	22	AA17957 Human TGF-beta rec
6	1748.5	73.6	800	22	AA18657 Human protein kina
7	1748.5	73.6	800	22	AA18657 Novel zipper conta
8	1409	59.3	349	22	AA175571 Human colon cancer
9	744	31.3	141	21	AA103563 Human secreted pro

10	638	26.9	124	22	AA186552
11	604	25.4	1021	23	AA1861000
12	601	25.3	1024	23	AA1864008
13	599.5	25.2	719	22	AA186513
14	599.5	25.2	1036	23	AA1860923
15	599.5	25.2	1036	23	AA1869779
16	599.5	25.2	1036	23	AA1862028
17	593	25.0	1036	23	AA1869648
18	590	24.8	1097	23	AA1862177
19	589	24.8	1097	23	AA1862177
20	581.5	24.5	847	23	AA1862263
21	568.5	23.9	859	24	AA1869745
22	565.5	23.8	859	24	AA1862886
23	565.5	23.8	859	18	AA1861227
24	565.5	23.8	859	24	AA1869744
25	564.5	23.8	888	23	AA1867049
26	544.5	22.9	1020	22	AA1865899
27	527.5	22.2	977	22	AA18671694
28	510.5	21.5	1490	22	AA18619123
29	504.5	21.2	850	23	AA186529
30	495	20.8	982	22	AA1860439
31	494	20.8	903	22	AA1860440
32	488	20.5	821	16	AA1860574
33	488	20.5	821	18	AA18617938
34	488	20.5	821	22	AA18650438
35	484	20.4	821	15	AA18646723
36	483	20.3	579	18	AA18627092
37	483	20.3	579	20	AA1862896
38	483	20.3	579	23	AA18685033
39	481	20.3	579	18	AA18627093
40	481	20.3	579	22	AA18609542
41	481	20.3	579	21	AA18691000
42	481	20.3	590	20	AA1869547
43	478.5	20.1	933	22	AA18650437
44	478.5	20.1	970	22	AA18650437
45	475.5	20.0	369	21	AA18622172

ALIGNMENTS

RESULT 1	ID	AA18657	standard; Protein; 455 AA.
AC	AA18657;		
XX			
DT	22-JAN-2001	(first entry)	
XX			
DE	A human regulator of intracellular phosphorylation.		
XX			
KW	Human; intracellular phosphorylation regulator; HRP; stroke; myeloma;		
KW	neurological disorder; Parkinson's disease; demyelinating disease;		
KW	meningitis; developmental disorder; neuromuscular disorder; cancer;		
KW	myasthenia gravis; cell proliferative disorder; actinic keratosis;		
KW	arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;		
KW	autoimmune disorder; inflammatory disorder; Addison's disease;		
KW	acquired immunodeficiency disease; allergy; diabetes mellitus;		
KW	rheumatoid arthritis; microbial infection; trauma.		
XX			
OS	Homo sapiens.		
XX			
FH	Domain	Location/Qualifiers	
FT	Modified-site	61	/note= "eukaryotic protein kinase domain"
FT	Modified-site	89	/note= "potential phosphorylation site"
FT	Modified-site	96	/note= "potential phosphorylation site"
FT	Modified-site	97	/note= "potential phosphorylation site"
FT	Modified-site	97	/note= "potential glycosylation site"

FT Binding-site 129..141
 FT /note= "protein kinase ATP-binding site"
 FT Modified-site 159
 FT /note= "potential glycosylation site"
 FT Modified-site 234
 FT /note= "potential phosphorylation site"
 FT Modified-site 252
 FT /note= "potential phosphorylation site"
 FT Modified-site 258
 FT /note= "potential phosphorylation site"
 FT Modified-site 265
 FT /note= "potential glycosylation site"
 FT Modified-site 268
 FT /note= "potential phosphorylation site"
 FT Region 294..322
 FT /note= "leucine zipper"
 FT Modified-site 302
 FT /note= "potential phosphorylation site"
 FT Modified-site 302
 FT /note= "potential phosphorylation site"
 FT Modified-site 342
 FT /note= "potential phosphorylation site"
 FT Modified-site 343
 FT /note= "potential phosphorylation site"
 FT Modified-site 346
 FT /note= "potential phosphorylation site"
 FT Modified-site 364
 FT /note= "potential phosphorylation site"
 FT Modified-site 409
 FT /note= "potential glycosylation site"
 FT Modified-site 410
 FT /note= "potential phosphorylation site"
 FT Modified-site 414
 FT /note= "potential phosphorylation site"
 FT Modified-site 415
 FT /note= "potential phosphorylation site"
 FT Modified-site 429
 FT /note= "potential phosphorylation site"
 FT Modified-site 434
 FT /note= "potential phosphorylation site"
 FT /note= "potential phosphorylation site"
 PN MO20005332-A2.
 XX
 PD 21-SEP-2000.
 PF 17-MAR-2000; 2000WO-US07277.
 XX
 PR 18-MAR-1999; 99US-0125593.
 PR 20-MAY-1999; 99US-0135049.
 PR 09-JUL-1999; 99US-0143188.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DM, Au-Young J;
 DR N-PSDB; AAA75674.
 DR WPI; 2000-602121/57.
 XX
 PT Novel human intracellular phosphorylation regulator polypeptides and
 PT polynucleotides for diagnosis, prevention and treatment of
 PT neurological, cell proliferative and autoimmune/inflammatory disorders
 PT
 XX
 PS Claim 1; Page 75-76; 96pp; English.
 XX
 CC The present sequence represents a human regulator of intracellular
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and

CC other developmental disorders of the central nervous system,
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
 CC rheumatoid arthritis, microbial infection and trauma.
 XX
 SQ Sequence 455 AA;
 Query Match 100.0%; Score 2375; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 5.3e-207;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLSGAPVQIKFDDLOFFENCSSFGSVYRAKISODKEVAVKKLKEKEAIIISVL 60
 DB 1 MSLSGAPVQIKFDDLOFFENCSSFGSVYRAKISODKEVAVKKLKEKEAIIISVL 60
 QY 61 SHRNIIQFYGVILPEPPNGIVTEYASLSGLYDYNINSSEEDMDHMTWADVAKGMHY 120
 DB 61 SHRNIIQFYGVILPEPPNGIVTEYASLSGLYDYNINSSEEDMDHMTWADVAKGMHY 120
 QY 121 LHMEDAPVKYIHRDLKSRNVVIADGVLLKICDFGASRFNHTTMSLVGTFPMMADEVIOQ 180
 DB 121 LHMEDAPVKYIHRDLKSRNVVIADGVLLKICDFGASRFNHTTMSLVGTFPMMADEVIOQ 180
 QY 181 LPSVETCDTYSYGVVLMWMLTREVPFKGLEQVAVLVKKERLLTIPSCPSFAELIH 240
 DB 181 LPSVETCDTYSYGVVLMWMLTREVPFKGLEQVAVLVKKERLLTIPSCPSFAELIH 240
 QY 241 QWEDADAKRPSFKIISTLESMSNDTSLPDKNSFLNKAARCEIATERLKLEBD 300
 DB 241 QWEDADAKRPSFKIISTLESMSNDTSLPDKNSFLNKAARCEIATERLKLEBD 300
 QY 301 LFFKEQLKERERRRLKMEQKLTQSNTPLLLPLAARMSSESYFSPSKTEESNSAEMSCOI 360
 DB 301 LFFKEQLKERERRRLKMEQKLTQSNTPLLLPLAARMSSESYFSPSKTEESNSAEMSCOI 360
 QY 361 TATSNGEGHGNPISLOAMLMGFQDI FSNKKAQAVHSGMOQINMOAKONSKTTSKRCK 420
 DB 361 TATSNGEGHGNPISLOAMLMGFQDI FSNKKAQAVHSGMOQINMOAKONSKTTSKRCK 420
 QY 421 KVNMLGFSDPDLSEGDGDDDDDDGEBDDMDNSE 455
 DB 421 KVNMLGFSDPDLSEGDGDDDDDDGEBDDMDNSE 455
 RESULT 2
 AA83278
 ID AAY83278 standard; Protein; 455 AA.
 XX
 AC AAY83278;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE Human survival regulating kinase (SRK).
 XX
 KW Survival regulating kinase; SRK; Raf; MAPK; BAD; MAPKK; MAPKK;
 KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
 KW apoptosis; cell survival; nuclear targeting; tumour; human;
 KW autoimmune disease.
 XX
 OS Homo sapiens.
 OS
 PN MO200022142-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 20-SEP-1999; 99WO-US22008.
 XX
 PR 13-OCT-1998; 98US-0104088.
 XX

PA (ONVX-) ONVX PHARM INC.
 XX Ruggieri R, Callow M, Diaz P;
 XX MPI; 2000-317994/27.
 DR N-PSDB; AA293783.
 XX
 PT Novel human survival regulating kinase polypeptide for screening agents
 PT which modulate biological pathways associated with SRK useful in
 PT treating autoimmune diseases, tumors and apoptosis-related disorders
 XX
 XX Claim 4; Figure 2; 62pp; English.
 CC Survival regulating kinases (SRK) are a class of proteins involved in
 CC cell signal transduction pathways such as mitogen-activated protein
 CC kinase pathways. A protein kinase activity means that the SRK can
 CC catalyze a reaction in which a phosphate group is transferred from a
 CC phosphate donor to a phosphate acceptor amino acid residue.
 CC preferably the hydroxyl side chain of a serine or threonine.
 CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
 CC activity is similar to that of a MAPKKK such as Raf. has a range of
 CC other activities including a cell growth-regulatory activity, a cell
 CC survival promoting activity, a HAX-1 binding activity, an apoptosis
 CC suppressing activity a MAPKK activation or stimulatory activity, a
 CC nuclear targeting activity and a SRK-specific immunogenic activity.
 CC SRK is useful for identifying agents which modulate cellular
 CC transformations mediated by Ras and SRK and agents that modulate the
 CC apoptosis suppression activity of SRK. This information may be useful
 CC in the treatment of autoimmune diseases, tumors and apoptosis
 CC related disorders.
 XX
 XX Sequence 455 AA;
 SQ
 Query Match 100.0%; Score 2375; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 5.3e-207;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 AC AAY84321;
 XX
 DT 12-JUL-2000 (first entry)
 XX
 DE A human cardiovascular system associated protein kinase-2.
 XX
 XX Human; cardiovascular system associated protein kinase-2; CSAPK-2;
 KW signaling pathway; cell growth; cell differentiation; gene mapping;
 KW tissue typing; forensic identification; cardiovascular disease;
 KW congestive heart failure; transgenic animal.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO200014212-A1.
 PN
 XX
 PD 16-MAR-2000.
 XX
 PF 09-SEP-1999; 99WO-US20631.
 XX
 XX 09-SEP-1998; 98US-009657.
 PR 29-SEP-1998; 98US-0163115.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Acton S;
 XX
 XX MPI; 2000-271053/23.
 DR N-PSDB; AA299726, AA299727.
 DR
 XX
 PT New nucleic acid encoding cardiovascular system associated protein
 PT kinase, used e.g. for diagnosis, treatment and prevention of
 PT cardiovascular disease
 XX
 XX Claim 2; Fig 2; 163pp; English.
 XX
 CC The present sequence represents a human cardiovascular system associated
 CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling
 CC pathways associated with cell growth and differentiation. The CSAPK
 CC polypeptides and polynucleotides are used to screen for agents that
 CC specifically modulate CSAPK, which are potential therapeutic agents.
 CC They are also used for diagnosis, prognosis or monitoring of
 CC CSAPK-related diseases, gene mapping, tissue typing and forensic
 CC identification, and for treating or preventing disorders associated
 CC with aberrant CSAPK expression or activity, especially cardiovascular
 CC diseases such as congestive heart failure. They can also be used in
 CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate
 CC transgenic animals.
 XX
 XX Sequence 455 AA;
 SQ
 Query Match 100.0%; Score 2375; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 5.3e-207;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 QCWEADAKRRPFFKQIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATELRLKLERD 300
 Qy 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
 Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
 Qy 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSRRGK 420
 Db 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSRRGK 420
 Qy 421 KVNMAIGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455
 Db 421 KVNMAIGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455
 RESULT 4
 AAM25322
 ID AAM25322 standard; Protein; 473 AA.
 XX
 AC AAM25322;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:837.
 XX
 KM Human; cancer; HIV infection; human immunodeficiency virus;
 KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiant; central nervous system; vitruide;
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KM antiagregant; haemostatic; vulnery; antilucer; osteopathic; eczema;
 KM dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
 KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antiapaplytic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.
 KM
 OS Homo sapiens.
 XX
 OS
 PN MO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000MO-US35017.
 XX
 PR 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI, 2001-457603/49.
 DR N-PSDB; AAH99263.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 Claim 20; Page 191; 1217p; English.
 XX
 PS AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antirhinitic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
 CC antilucer; osteopathic; dermatological; antiallergic; antiaesthetic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 CC
 XX
 SQ Sequence 473 AA;
 Query Match 99.6%; Score 2366; DB 22; Length 473;
 Best Local Similarity 99.8%; Pred. No. 3,7e-206;
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MSLSGASFWQIKFPDLOFPENCGGSPGSGVYAKMTISQKXAVKLLKIEKEAETLSVL 60
 Db 19 MSSLGASFWQIKFPDLOFPENCGGSPGSGVYAKMTISQKXAVKLLKIEKEAETLSVL 78
 Qy 61 SHRNIQFYGVILPEPNYGVITEYASIGSLYDINSNRSEMDMHIIMTATDVAKGMHY 120
 Db 79 SHRNIQFYGVILPEPNYGVITEYASIGSLYDINSNRSEMDMHIIMTATDVAKGMHY 138
 Qy 121 LHMEAPVKVIHRLDKSRNVVIAADGVLCIDFGASRFHNTTHMSIVGTFPMMAPEVIO 180
 Db 139 LHMEAPVKVIHRLDKSRNVVIAADGVLCIDFGASRFHNTTHMSIVGTFPMMAPEVIO 198
 Qy 181 LPVSEFCPTYGVIVLWMLTREVPKLEGIOVAMLVYKNERLTPSCCRSPFELH 240
 Db 199 LPVSEFCPTYGVIVLWMLTREVPKLEGIOVAMLVYKNERLTPSCCRSPFELH 258
 Qy 241 QCWEADAKRRPFFKQIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATELRLKLERD 300
 Db 259 QCWEADAKRRPFFKQIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATELRLKLERD 318
 Qy 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
 Db 319 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 378
 Qy 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSRRGK 420
 Db 379 TATSNBEGHGMNPSLOAMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSRRGK 438
 Qy 421 KVNMAIGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455
 Db 439 KVNMAIGFSDPDLSEGDGDDDDGEEEDNDMDNSE 473
 RESULT 5
 AAB71957
 ID AAB71957 standard; Protein; 800 AA.
 XX
 AC AAB71957;
 XX
 DT 11-MAY-2001 (first entry)
 XX
 DE Human TGF-beta receptor encoded by cdna clone HDP5M48.
 XX
 KM Human; antisclerotic; dermatological; immunosuppressive; cytostatic;
 KM antiinflammatory; anti-HIV; immunostimulant; cardiant; vascular;
 KM ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;
 KM antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;
 KM transforming growth factor; TGF-beta receptor; immune disorder;
 KM hyperproliferative disorder; cardiovascular disease; angiogenesis;
 KM neurological disorder.
 KM
 OS Homo sapiens.
 XX

QY 1 MSLSGASFWQIKFDDLOFPENCSSGSGSVYRAKWIISQDEKAVAVKKLTKIEKAEILSVL 60
 DB 1 MSLSGASFWQIKFDDLOFPENCSSGSGSVYRAKWIISQDEKAVAVKKLTKIEKAEILSVL 60
 QY 61 SHRNIIQFYGVILPEPNYGVITEVASLGSLYDYINSNSREEMDMHMTWATDVAKGMY 120
 DB 61 SHRNIIQFYGVILPEPNYGVITEVASLGSLYDYINSNSREEMDMHMTWATDVAKGMY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIADGVKICDPGASRFHNHTTHMSLVGTFFPMMAPEVIO 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIADGVKICDPGASRFHNHTTHMSLVGTFFPMMAPEVIO 180
 QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVYKERNLTTPSCCRSPFALH 240
 DB 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVYKERNLTTPSCCRSPFALH 240
 QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSEFLHNKAERCEIATELTKLTERD 300
 DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSEFLHNKAERCEIATELTKLTERD 300
 QY 301 LSPFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355
 DB 301 LSPFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355
 QY 356 MSCQITATNSGEGHGMNPSLOAMMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404
 DB 356 MSCQITATNSGEGHGMNPSLOAMMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404
 QY 405 ---QAKONSSKTTSKRGKKVMALGFSDFLSEGDHDDDD-----DDGEE 446
 DB 405 ---QAKONSSKTTSKRGKKVMALGFSDFLSEGDHDDDD-----DDGEE 446
 QY 416 HFPPILKDSGGEPEENBEKIVLLELVFG-FHLKPGTGPDCKKMYMENDGDE 467
 DB 416 HFPPILKDSGGEPEENBEKIVLLELVFG-FHLKPGTGPDCKKMYMENDGDE 467

RESULT 7
 ABP43736
 ID ABP43736 standard; Protein; 800 AA.
 AC ABP43736;
 XX
 XX 26-FEB-2003 (first entry)
 DE Motif zipper containing kinase AZK.
 XX
 XX Neuroprotective; immunomodulator; cancer; chromosome 2q24.2;
 KW cytostatic; anti-inflammatory; gene therapy; nutritional supplement;
 KW wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KW vulnertary.
 XX
 XX Homo sapiens.
 OS
 PN WO200231111-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 11-OCT-2001; 2001MO-US27760.
 XX
 PR 12-OCT-2000; 2000US-0687527.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Aundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 DR WPI; 2002-426278/45.
 DR N-PSDB; ABQ60980.
 XX
 PT New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation

PS Claim 20; SEQ ID # 639; 357bp + sequence listing; English.
 XX
 CC The invention relates to 446 newly isolated polynucleotide sequences.
 CC The activity of polynucleotides of the invention may be described as,
 CC vulnertary, neuroprotective, immunomodulator, cyostatic and
 CC anti-inflammatory. Compositions comprising nucleic acids of the invention
 CC are useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records
 CC ABP43544-ABP43989 represent polypeptides encoded by polynucleotides of
 CC the invention.
 CC NOTE: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 800 AA;

Query Match 73.6%; Score 1748.5; DB 23; Length 800;
 Best Local Similarity 76.1%; Pred. No. 1.1e-149;
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSLSGASFWQIKFDDLOFPENCSSGSGSVYRAKWIISQDEKAVAVKKLTKIEKAEILSVL 60
 DB 1 MSLSGASFWQIKFDDLOFPENCSSGSGSVYRAKWIISQDEKAVAVKKLTKIEKAEILSVL 60
 QY 61 SHRNIIQFYGVILPEPNYGVITEVASLGSLYDYINSNSREEMDMHMTWATDVAKGMY 120
 DB 61 SHRNIIQFYGVILPEPNYGVITEVASLGSLYDYINSNSREEMDMHMTWATDVAKGMY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIADGVKICDPGASRFHNHTTHMSLVGTFFPMMAPEVIO 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIADGVKICDPGASRFHNHTTHMSLVGTFFPMMAPEVIO 180
 QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVYKERNLTTPSCCRSPFALH 240
 DB 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVYKERNLTTPSCCRSPFALH 240
 QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSEFLHNKAERCEIATELTKLTERD 300
 DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSEFLHNKAERCEIATELTKLTERD 300
 QY 301 LSPFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355
 DB 301 LSPFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355
 QY 356 MSCQITATNSGEGHGMNPSLOAMMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404
 DB 356 MSCQITATNSGEGHGMNPSLOAMMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404
 QY 405 ---QAKONSSKTTSKRGKKVMALGFSDFLSEGDHDDDD-----DDGEE 446
 DB 405 ---QAKONSSKTTSKRGKKVMALGFSDFLSEGDHDDDD-----DDGEE 446
 QY 416 HFPPILKDSGGEPEENBEKIVLLELVFG-FHLKPGTGPDCKKMYMENDGDE 467
 DB 416 HFPPILKDSGGEPEENBEKIVLLELVFG-FHLKPGTGPDCKKMYMENDGDE 467

RESULT 8
 AAG75571
 ID AAG75571 standard; Protein; 349 AA.
 AC AAG75571;
 XX
 XX 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:6335.
 XX
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.

XX 05-APR-2001.
 PD 28-SEP-2000; 2000WO-US26524.
 XX
 PF 29-SEP-1999; 99US-0157137.
 XX PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-235357/24.
 DR N-PSDB; AAH34976.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 7789-7790; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing P.
 CC Inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 XX Sequence 349 AA;
 SQ
 Query Match 59.3%; Score 1409; DB 22; Length 349;
 Best Local Similarity 100.0%; Pred. No. 2.4e-119;
 Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSLGASFOVKEDLDQFENCGSGSGSVYRAKWTISODKEVAVKLLKTEKEAEILSVL 60
 Db MSSLGASFOVKEDLDQFENCGSGSGSVYRAKWTISODKEVAVKLLKTEKEAEILSVL 116
 QY 61 SHRNIIQFVGVIIEPPNYGIVTEYASIGSLYDIYNSRSEMDMHIIMTAVTAVAKMHY 120
 Db SHRNIIQFVGVIIEPPNYGIVTEYASIGSLYDIYNSRSEMDMHIIMTAVTAVAKMHY 176
 QY 121 LHMEAPVKVIHRDLKSRNVIIADGVLTICDPGASRFPHNHTTMSLVGTPEMMAPEVIO 180
 Db LHMEAPVKVIHRDLKSRNVIIADGVLTICDPGASRFPHNHTTMSLVGTPEMMAPEVIO 226
 QY 181 LPVSETCDITYSYGVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTIPSSCPRSFALLH 240
 Db LPVSETCDITYSYGVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTIPSSCPRSFALLH 256
 QY 241 QCWEADAKKRPSTFKQIISILESMSNDTSL 269
 Db 297 QCWEADAKKRPSTFKQIISILESMSNDTSL 325
 RESULT 9
 AAG03583
 ID AAG03583 standard; Protein; 141 AA.
 XX
 AC AAG03583;
 XX

DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 7664.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Malne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR N-PSDB; AAC03589.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 7664; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 XX Sequence 141 AA;
 SQ
 Query Match 31.3%; Score 744; DB 21; Length 141;
 Best Local Similarity 98.6%; Pred. No. 1.7e-59;
 Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 102 MDMDHIMTAVTAVAKMHYLAHPVKVIHRDLKSRNVIIADGVLTICDPGASRFPHNHT 161
 Db 1 MDMDHIMTAVTAVAKMHYLAHPVKVIHRDLKSRNVIIADGVLTICDPGASRFPHNHT 60
 QY 162 TMSLVGTPEMMAPEVIOGLPVSETCDITYSYGVLMEMLTREVPEFKGLEGLQVAMLVVEK 221
 Db TMSLVGTPEMMAPEVIOGLPVSETCDITYSYGVLMEMLTREVPEFKGLEGLQVAMLVVEK 120
 QY 222 NERLTIPSSCPRSFALLHQC 242
 Db 121 NERLTIPSSCPRSFALLHQC 141
 RESULT 10
 AAB65552
 ID AAB65552 standard; Protein; 124 AA.
 XX
 AC AAB65552;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE C-terminus specific to novel human protein kinase MLK4B.
 XX

KM Human; protein kinase; antiarthritic; antisclerotic; immunosuppressive;
 KM Cardiac; renal; antiinflammatory; antiaslomatic; osteopathic;
 KM dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KM immune disorder; cardiovascular disease; neurodegenerative disease;
 KM cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KM inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX Homo sapiens.
 OS
 XX
 XX WO200073469-A2.
 XX
 XX 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000MO-US14842.
 XX
 XX 28-MAY-1999; 99US-0136503.
 XX
 XX (SUGB-) SUGEN INC.
 XX
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 PI
 XX WPI; 2001-032161/04.
 XX
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 XX Example 1; Page 116; 310pp; English.
 XX
 XX The present sequence is given in a specification relating to novel
 CC protein kinases. The protein kinases and the nucleic acids that encode
 CC them may be used in the treatment and diagnosis of diseases
 CC associated with inappropriate kinase expression such as immune-related
 CC diseases and/or cancers. The nucleic acids and complementary sequences
 CC may also be used as DNA probes in diagnostic assays. The kinase
 CC polypeptides may be used as antigens in the production of antibodies of
 CC kinase expression and activity. Anti-kinase antibodies and kinase
 CC antagonists may also be used to down regulate kinase expression and
 CC activity. Diseases related to kinase expression and activity include
 CC rheumatoid arthritis, atherosclerosis, autoimmune disorders,
 CC complications of organ transplantation, myocardial infarction, immune
 CC disorders, cardiomyopathies, strokes, renal failure, oxidative-stress
 CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
 CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
 CC disorders.
 CC
 XX
 XX Sequence 124 AA;
 SQ
 XX
 XX Query Match 26.9%; Score 638; DB 22; Length 124;
 Best Local Similarity 100.0%; Pred. No. 6.1e-50;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 332 LPLAAMSESEYESTKEESNAEMSCQITATNGEGHGNPGLQAMMLMGFDISMK 391
 DB 1 LPLAARSESEYESTKEESNAEMSCQITATNGEGHGNPGLQAMMLMGFDISMK 60
 QY 392 AGAVMHSQMIOAKONSSKTSKRGKLVNNAIGFSDPLESGDDDDDDDEEDNDM 451
 DB 61 AGAVMHSQMIOAKONSSKTSKRGKLVNNAIGFSDPLESGDDDDDDDEEDNDM 120
 QY 452 DNSE 455
 DB 121 DNSE 124

DT 10-SEP-2002 (first entry)
 XX
 XX DE Novel human protein. SEQ ID 87.
 XX
 XX Human; cytostatic; vulnary; antiarteriosclerotic; antiParkinsonian;
 KM neurotropic; neuroprotective; immunosuppressive; haemostatic;
 KM antiinflammatory; cardiac; antitumor; antitumor; antitumor;
 KM cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 KM wound healing disorders; atherosclerosis; Parkinson's disease;
 KM Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 KM inflammation; neoplastic disease; nervous system disorder;
 KM cardiovascular disorders; systemic autoimmune disease; hyperimmunity;
 KM hyperproliferation; systemic autoimmune disease; hyperimmunity;
 KM developmental abnormality; gastrointestinal ulceration; neuropathy;
 KM haematological disease; metabolic disease; sperm dysfunction;
 KM thyroid disorder; hypothyroidism; brain damage; colitis;
 KM cone photo-transduction deficiency; neurological disease; stroke;
 KM angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 KM trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KM growth abnormality; precocious puberty.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200250105-A1.
 XX
 XX 27-JUN-2002.
 XX
 XX 17-DEC-2001; 2001WO-US49232.
 XX
 XX 19-DEC-2000; 2000US-256710P.
 XX 20-DEC-2000; 2000US-257048P.
 XX 09-JAN-2001; 2001US-260482P.
 XX 30-JAN-2001; 2001US-264922P.
 XX 06-FEB-2001; 2001US-266797P.
 XX 19-MAR-2001; 2001US-276988P.
 XX 04-APR-2001; 2001US-281535P.
 XX 08-MAY-2001; 2001US-289622P.
 XX
 XX (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 XX Agarwal P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
 PI Marcensen SA, Ritzvi SK, Smith RF, Strum JC, Xie Q;
 XX WPI; 2002-508784/54.
 DR N-PSDB; ABQ86165.
 XX
 XX Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune
 PT disorder -
 XX
 XX Claim 1(a); Page 307-309; 335pp; English.
 XX
 XX The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated.
 CC The activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnary, antiarteriosclerotic, antiParkinsonian, neurotropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiac, antitumor, antitumor, antitumor, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder.
 CC hyperproliferation, systemic autoimmune disease, hyperimmunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,

CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABB60965-ABB61019 represent novel human
 CC proteins of the invention.

XX Sequence 1021 AA;

Query Match 25.4%; Score 604; DB 23; Length 1021;
 Best Local Similarity 34.4%; Pred. No. 1.8e-45;
 Matches 151; Conservative 74; Mismatches 142; Indels 72; Gaps 10;

QY 6 ASFVQIKFDLPDLPFENCGGSGFSGVYRAKMSODKEVAVKLIK-----TEKE 53
 DB 114 SSVFVHVAFERLEKEILGAGGFGQVVRATW--QCGEVAVAAQODEQDAAAASVRR 171
 QY 54 AELLSVLSHRNIIQFPGVILPEPNYGIWTEYASLSGLYDINSRSEMDMDHIM-TWAT 112
 DB 172 ARLEFAMLRHFNITELRGVCLQCPHLCLVLEFARGALNRALARRIP---PHVLVWAV 227
 QY 113 DVAKGHHVLMKAPVVIHRDLKSRNVV-----AADGVLKICDFGASRFHNTTHM 164
 DB 228 QIARGMLVHLEBAFVPLIHRDLKSNILLLEKTEHDDICNKTKITDFGLAREWHRTTK 287
 QY 165 SLVGTFFPMAPREVIOSLPVSETCTYSYGVLMEMLTREVPFKLEGLOVAMLVKNER 224
 DB 288 STAGTAMAPKPEYIKSLPSKSDISYGLTWELLTGVPYRGIDGLAVAYAVANKLT 347
 QY 225 LTPSSCPSPFAELHOCWEADAKKRPSEKQIISIESMSND--TSLPDKNSFLHNKAE 282
 DB 348 LRPSPCEPPEFALMKCEWQDPHRRPSEFALLLEQTLAIEGAVWTEPQF--SFHSMQD 405
 QY 283 WRCEIATLERLKLEKRLDSFKQEL-----KERERLKMEQKLTREQ----- 325
 DB 406 WKLEIQMDELTKERELKSRSEELRRAALQOKSOBELKRRQOLAEEDIVLEREIN 465
 QY 326 -----SNTPLLLPLAARMESESYFESKTEESNSAEM-----SCQIT 361
 DB 466 ILIFQINQEKPKYKRRKKRKRRLTKDKGHRISLSPDPQHKITTVQASPLMDRRRLNS 525
 QY 362 AITSNGEGHGMNPSLOAMML 380
 DB 526 SSSPPSSPTMMPRLRAIQL 544

RESULT 12
 ABB98408 standard; Protein; 1024 AA.
 XX ABB98408;
 AC
 XX
 XX 21-OCT-2002 (first entry)
 DT
 XX
 DE Human NOV7, a mixed lineage kinase 2-like protein.
 XX
 XX Human: NOV7, cytosolic; Cardiant; Antiinflammatory; Immunosuppressive;
 KM Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
 KM Antistimatic; Nephrotropic; Hepatotropic; Neutroprotective; Nootropic;
 KM Antibacterial; Vitruclide; Antiparasitic; Relaxant; Anticonvulsant;
 KM Gene Therapy; NOV; cancer; heart disease; inflammation;
 KM autoimmune disorder; allergy; blood disorder; AIDS; diabetes;
 KM obesity; asthma; IGA nephropathy; cirrhosis; arthritis;
 KM Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
 KM wasting disorder; mixed lineage kinase 2-like protein.
 KM
 XX Homo sapiens.
 OS
 XX
 XX MO200255704-A2.
 PN
 XX 18-JUL-2002.
 PD
 XX 09-JAN-2002; 2002WO-US00554.
 PF
 XX 09-JAN-2001; 2001US-260417P.
 PR

PR 10-JAN-2001; 2001US-260831P.
 PR 28-FEB-2001; 2001US-272338P.
 PR 09-MAR-2001; 2001US-274876P.
 PR 18-APR-2001; 2001US-284704P.

XX (CURA-) CURAGEN CORP.

XX Padigar M, Li L, Zernusen BD, Casman SJ, Shenoy S, Spytek KA,
 PI Zhong M, Gangoli EA, Burgess CE, Paturajan M, Verneet CAM,
 PI Taylor S, Tchernov VT, Miller CE, Guo X, Boidog FL, Grosse WM,
 PI Alsdorok JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K,
 PI MacDougall J, Maityankar U, Miller I, Peyman J, Smithson G,
 PI Gunther E, Stone DJ,
 XX
 XX WPI, 2002-590674/63.
 DR N-PSDB; ABB85385.

PT NOVX polypeptides and encoding polynucleotides, useful for preventing
 PT or treating NOVX-associated disorders e.g. cancer, inflammation, or
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 PT pharmacogenomics -
 XX
 XX Claim 1; Page 57; 358pp; English.

CC The present sequence is the protein sequence for a NOV protein. The
 CC NOV proteins and coding sequences are useful for treating or preventing
 CC NOV-associated disorders or in the manufacture of a medicament for
 CC treating the disorders, such as cancer, heart disease, inflammation,
 CC autoimmune disorders, allergies, blood disorders, AIDS, diabetes,
 CC obesity, asthma, IGA nephropathy, cirrhosis, arthritis, Alzheimer's
 CC disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular
 CC dystrophy, epilepsy, and other wasting disorders associated with chronic
 CC diseases. NOV7 is a mixed lineage kinase 2-like protein, and the NOV7
 CC coding sequence is localised to chromosome 14q24.3.
 CC
 XX

SQ Sequence 1024 AA;

Query Match 25.3%; Score 601; DB 23; Length 1024;
 Best Local Similarity 39.7%; Pred. No. 3.5e-45;
 Matches 136; Conservative 58; Mismatches 117; Indels 32; Gaps 5;

QY 10 QIKFDLPDLPFENCGGSGFSGVYRAKMSODKEVAVK-----KLKTEKEAEIT 57
 DB 126 EIDFAELTLEBIIGIGFGKVTAEWIGD--EVAVAARHPDDEDISQTIENRQBAKLF 183
 QY 58 SVLSHRNIIQFPGVILPEPNYGIWTEYASLSGLYDINSRSEMDMDHIMTWATDVAK 117
 DB 184 AMLKHPNIIALRGVCLKEPNLCVMEFARGGPLNRLSGKR---IPDLVMAVQIANG 240
 QY 118 MHYLMKAPVVIHRDLKSRNVV-----AADGVLKICDFGASRFHNTTHMSLVGT 169
 DB 241 MNYLHDEAVIPIIHRDLKSNVLIQKVENGDLSNKLITITDFGLAREWHRTTKSAAGT 300
 QY 170 FPMMAPEVIOSLPVSETCTYSYGVLMEMLTREVPFKLEGLOVAMLVKNERRTTSS 229
 DB 301 YAMMAPEVIRASNFSGSVWSYGLTWELLTGEVFPFGIDGLAVAYAVANKLAPITS 360
 QY 230 SCPRSPFAELHOCWEADAKKRPSEKQIISIESMSNDTSLPDKNSFLHNKAEWRCEIA 289
 DB 361 TCCEPAPKLMCEOWNDPDRSRPSFTIIDLQTLTIESGFEMPKSFHCLQDNWKEIJD 420
 QY 290 TLERLKKRLDSFKQEL-----KERERLKMEQKLTREQ 325
 DB 421 MFDOLRAKEKEELRWEEELTRALQOKQDEELRRREDELAER 463

RESULT 13
 AAB85513
 ID AAB85513 standard; protein; 719 AA.
 XX
 XX AAB85513;
 AC
 XX
 XX 25-SEP-2001 (first entry)
 DT

DE Human protein kinase SGK067.

KM Protein kinase; enzyme; cytosolic; nootropic; neuroprotective; human;

KM antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;

KM analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic;

KM antipsoriatic; antihemematic; antiarthritic; ophthalmological; anorectic;

KM osteopathic; thrombolytic; antiarteriosclerotic; antislathmatic;

KW vasootropic; antidiabetic; gene therapy.

OS Homo sapiens.

XX WO200155356-A2.

XN 02-AUG-2001.

PD 25-JAN-2001; 2001WO-US02337.

PF 25-JAN-2001; 2000US-0178078.

XX 25-JAN-2000; 2000US-0178078.

PR 31-JAN-2000; 2000US-0179364.

PR 17-FEB-2000; 2000US-0183173.

PR 17-MAR-2000; 2000US-0190162.

PR 29-MAR-2000; 2000US-0193404.

PR 13-NOV-2000; 2000US-0247013.

XX (SUGG-) SUGEN INC.

PA Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;

P1 WPI; 2001-476202/51.

XX N-PDB; AAH46913.

DR

DR

XX

XX

PS Claim 7, Page 217, 218pp; English.

XX The invention provides human protein kinases and protein kinase-like
enzymes and polynucleotides encoding the polypeptides. The kinase
CC polypeptides and their modulators are useful for treating a disease or
disorder such as cancer, immune-related diseases, cardiovascular disease,
brain or neuronal-associated disease and metabolic disorders, including
cancers of tissues, cancers of hematopoietic origin, diseases of the
central nervous system, diseases of the peripheral nervous system,
Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
lateral sclerosis, viral infections, infections caused by prions,
bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,
mood disorders, attention disorders, cognition disorders, hypotension,
hypertension, psychotic disorders, neurological disorders, dyskinesias,
metabolic disorders, and organ transplant rejection. They are also useful
for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,
osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic
inflammatory bowel disease, rheumatoid arthritis, metabolic disorders
such as diabetes, obesity, cardiovascular diseases such as reperfusion
injury, coronary thrombosis, clotting disorders and atherosclerosis,
ocular diseases such as glaucoma, retinopathy and macular degeneration,
psychiatric and neurological disorders such as anxiety, schizophrenia,
dementia, manic depression, etc. The polynucleotides are useful in gene
therapy techniques to treat the above mentioned disorders. Sequences
AA85491-85522 represent the human protein kinases of the invention.

XX

SQ Sequence 719 AA;

Query Match 25.2%; Score 599.5; DB 22; Length 719;
Best Local Similarity 33.3%; Pred. No. 2.8e-45;
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9

OY 6 ASFOVKRDDIQLFFPNCGGSGFSGYSRKWKISQDKEVAVKKLK-----IEKE 53
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 114 SSPHVNAERLEKRLIGAGCGCYRRATW--QGGEVAVKARKQRPEDMAAASVRE 171
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

YY 54 AEIIIVLSHRNIIOFYGVILPEPNYGVITEVASLGSLYDI-----NSNRSEE 101

Db	172	ARLFAMLEHPNIIELRGVCLQOPHLCVLEFARGALURALAANAAPDPRAAPRRARR	231
Qy	102	MDMDHIMTWADVAKGMHYLMNEAPVKYIHRDLKSRNVI-----AADGVLCICDG	153
Db	232	IPPHLVNMAVOIARGMLYLHEEAFVPIIHRDLKSSNIILEKIEHDDICNKTLCIDFG	291
Qy	154	ASRFPHNHTHSLSVGFPEPMNAPEVQSLLPVSEPCITYSGVGLWEMLTREVPFKGLEGLQ	213
Db	292	LAREHRTTKTSTACTAYAMNAPEVKSLSFKGSDIWSYGVLLMELLTGEVYRGIDGLA	351
Qy	214	VAMLVENKEBELTTPSSCPSPFAELHOCWEADAKKRPSCFOIISLESMSND--TSLPD	271
Db	352	VAYGAVNKKLTLPISTCPPEPAKLMKSCWQDDPHIRPSFALILEQLTAIEGAVMTMPQ	411
Qy	272	KNSGLHNKAEMRCIEATLEBKLEEDLSKQDEL-----KERERLKMWEKLTLE	324
Db	412	E-SFHSQDDMKUEIQCMFDELRTKEKELRSEBELTRAAQQOSQEBLLKRREQLAE	469
Qy	325	Q-----SNTPLLLPLARMSESYFESKTEESNSA	354
Db	470	REIDVLERELNILLFQLNOEKPKVKRKGKFKRSRLKLDGHRISLPSPDFHKITVQASP	529
Qy	355	EM---SCQITATSGEGHGMNPSLQAMML	380
Db	530	NLDKRRSLNSSSSPSPSSPTMPLRALQL	559
RESULT 14			
ID	ABB80923	standard; Protein; 1036 AA.	
AC	ABB80923;		
XX			
DT	08-OCT-2002	(first entry)	
XX			
DE	Novel human protein (NHP) kinase.		
KM	Novel human protein; NHP; kinase; human; enzyme.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	Misc-difference 925	/note= "encoded by WGT"	
XX			
PN	WO200255685-A2.		
PD	18-JUL-2002.		
XX			
PF	10-DEC-2001; 2001WO-US47606.		
XX			
PR	11-DEC-2000; 2000US-254744P.		
XX			
PA	(LEXI-) LEXICON GENETICS INC.		
PI	Hu Y, Kieke JA, Donoho G;		
XX			
DR	WPI; 2002-566739/60.		
XX			
PT	N-PSDB; ABN86357, ABN86358.		
XX			
PS	Novel human kinase polynucleotide encoding a protein that shares structural similarity with animal kinases for therapeutic, diagnostic and pharmacogenomic applications -		
XX			
CC	Claim 1; Page 37-39; 41pp; English.		
CC	The invention relates to a novel human protein (NHP), kinase that shares structural similarity with animal kinases. The kinase polynucleotides are useful in therapeutic, diagnostic and pharmacogenomic applications and for identifying compounds that modulate, i.e. act as agonists or antagonists of the gene expression or gene product activity. The present sequence represents the NHP kinase.		

XX SQ Sequence 1036 AA;
 Query Match 25.2%; Score 599.5; DB 23; Length 1036;
 Best Local Similarity 33.3%; Pred. No. 4,8e-45;
 Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;
 QY 6 ASFVQIKFPDLOPFENGCGSGSVYRAKMIISQDKXVAVKLLK-----IEKE 53
 DB 114 SFPVAVAFERLELKEKELIGAGFGQYRATW--QGQEVAVKARQDEQDAAAAAASVRE 171
 QY 54 AEILSVLSHRNTIQFYGVILEPPNYGIYVEYASLSGLYDI-----NSNRSEE 101
 DB 172 ARLFAMLRHPNITIELRGVCLQOPHLCLVLEFARGALNRLAANAAPPDPRAPRRARR 231
 QY 102 MDMDHIMTATDVAKGMHYLHMEAPYKVIHRDLKSNVYI-----AADGVLTICDFG 153
 DB 232 IPPHYLVNMAVOIARGMVLHBEAFVPIIHRDLKSSNILLKEIHDDICNKTLTITDFG 291
 QY 154 ASRFHHTHMSLVGFPPMAAPRVISLPSVSECTDYSGVVLWEMLTREVPFKLEGHQ 213
 DB 292 LAREHRTTKMSTAGTYAMAPBEVILKSLFSKGSIDWSYGVLLWELTGBVPRGIDGLA 351
 QY 214 VAMLVKXNERLTIIPSSCPSPFAELLHQCEWADAKRPSFKQIISLESMSND--TSLPD 271
 DB 352 VAYGAVVNKLTLPIPTCEPFAKLMKECQODPHIRPFAILLBQTLIEGAVTTEMPQ 411
 QY 272 KGNSEFLHAKAEMRCIEATLERLKLRLDSFKEQEL-----KERERLKMWEQKLTE 324
 DB 412 E--SFHSMODDWKLEIQMFDELRTKEKELRSREELTRALQOKSQEELLKRREQOLAE 469
 QY 325 Q-----SNTPLLPLAARMSESYFESKTEESNQA 354
 DB 470 REIDVLERELNLIIFOLNOEKPRVKRKGKFKRSRLKLDGHRISLPSDFQHKITYQASP 529
 QY 355 EM-----SCQITATSNCEGHGMNPSLOAMML 380
 DB 530 NLDKRSLSNSSSSPSSPTMMRLRAIQL 559
 RESULT 15
 ABB99779
 ID ABB99779 standard; Protein; 1036 AA.
 AC ABB99779;
 DT 24-MAR-2003 (first entry)
 XX
 DE Amino acid sequence of human mixed lineage kinase 7 (MLK7).
 KW Human; mixed lineage kinase 7; MLK7; chromosome 1; cancer;
 XX neurodegenerative disorder; inflammation.
 OS Homo sapiens.
 XX
 OS
 PN WO200295017-A1.
 PD 28-NOV-2002.
 XX
 PF 23-MAY-2002; 2002WO-US16387.
 XX
 PR 24-MAY-2001; 2001US-293381P.
 XX
 PA (CEPH-) CEPHALON INC.
 PI Angeles TS, Durkin JT, Holakin BP, Meyer SL, Spals CM;
 XX MPI; 2003-148466/14.
 DR N-PSDB; AB223269.
 XX
 PT New isolated MLK7 polynucleotide and polypeptide, useful for the
 PT diagnosis or treatment of disorders with aberrant expression or
 PT activity of the MLK7 polypeptide, such as cancer, neurodegenerative

PT disorders and inflammation -
 XX
 PS Claim 12; Page 41-44; 68pp; English.
 CC The present sequence represents a human mixed lineage kinase 7 (MLK7)
 CC polypeptide. The MLK7 gene maps to location q42.2-q43 of chromosome 1.
 CC The MLK7 polypeptide contains a leucine zipper domain. MLK7 polypeptides
 CC and polynucleotides are useful for the diagnosis and treatment of
 CC disorders associated with the aberrant expression or activity of MLK7,
 CC such as cancer, neurodegenerative disorders and inflammation.
 XX
 SQ Sequence 1036 AA;
 Query Match 25.2%; Score 599.5; DB 24; Length 1036;
 Best Local Similarity 33.3%; Pred. No. 4,8e-45;
 Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;
 QY 6 ASFVQIKFPDLOPFENGCGSGSVYRAKMIISQDKXVAVKLLK-----IEKE 53
 DB 114 SFPVAVAFERLELKEKELIGAGFGQYRATW--QGQEVAVKARQDEQDAAAAAASVRE 171
 QY 54 AEILSVLSHRNTIQFYGVILEPPNYGIYVEYASLSGLYDI-----NSNRSEE 101
 DB 172 ARLFAMLRHPNITIELRGVCLQOPHLCLVLEFARGALNRLAANAAPPDPRAPRRARR 231
 QY 102 MDMDHIMTATDVAKGMHYLHMEAPYKVIHRDLKSNVYI-----AADGVLTICDFG 153
 DB 232 IPPHYLVNMAVOIARGMVLHBEAFVPIIHRDLKSSNILLKEIHDDICNKTLTITDFG 291
 QY 154 ASRFHHTHMSLVGFPPMAAPBEVISOPLVSECTDYSGVVLWEMLTREVPFKLEGHQ 213
 DB 292 LAREHRTTKMSTAGTYAMAPBEVILKSLFSKGSIDWSYGVLLWELTGBVPRGIDGLA 351
 QY 214 VAMLVKXNERLTIIPSSCPSPFAELLHQCEWADAKRPSFKQIISLESMSND--TSLPD 271
 DB 352 VAYGAVVNKLTLPIPTCEPFAKLMKECQODPHIRPFAILLBQTLIEGAVTTEMPQ 411
 QY 272 KGNSEFLHAKAEMRCIEATLERLKLRLDSFKEQEL-----KERERLKMWEQKLTE 324
 DB 412 E--SFHSMODDWKLEIQMFDELRTKEKELRSREELTRALQOKSQEELLKRREQOLAE 469
 QY 325 Q-----SNTPLLPLAARMSESYFESKTEESNQA 354
 DB 470 REIDVLERELNLIIFOLNOEKPRVKRKGKFKRSRLKLDGHRISLPSDFQHKITYQASP 529
 QY 355 EM-----SCQITATSNCEGHGMNPSLOAMML 380
 DB 530 NLDKRSLSNSSSSPSSPTMMRLRAIQL 559
 Search completed: December 5, 2003, 09:18:23
 Job time : 54.2593 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: December 5, 2003, 09:20:22 ; Search time 35.6481 Seconds
(without alignments)
2373.828 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375
Sequence: 1 MSLSGASFYQIKFDLDQFPE.....GDDDDDDDEEDNDNDNSE 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2375	100.0	455	9	US-09-757-982-5
2	2375	100.0	455	12	US-10-094-749-2477
3	1409	59.3	349	15	US-10-106-698-6345
4	599.5	25.2	1036	12	US-10-354-358-24
5	599.5	25.2	1036	14	US-10-014-882-2
6	593	25.0	394	10	US-09-862-027-19
7	590	24.8	1097	12	US-10-288-798-12
8	581.5	24.5	847	14	US-10-143-133-2
9	568.5	23.9	966	15	US-09-771-161A-197
10	565.5	23.8	746	15	US-10-153-668-436
11	565.5	23.8	859	15	US-10-153-668-324
12	565.5	23.8	892	15	US-10-153-668-438
13	545.5	23.0	328	10	US-09-862-027-18
14	504.5	21.2	850	10	US-09-904-389-2
15	488	20.5	821	12	US-10-171-404A-48

16	487	20.5	263	10	US-09-840-704-5	Sequence 5, Appl1
17	481	20.3	579	12	US-10-384-743-4	Sequence 4, Appl1
18	481	20.3	579	14	US-10-158-895-4	Sequence 4, Appl1
19	481	20.3	590	12	US-10-384-743-15	Sequence 15, Appl1
20	481	20.3	590	14	US-10-158-895-15	Sequence 15, Appl1
21	471.5	19.9	518	15	US-10-283-023-2	Sequence 2, Appl1
22	424	17.9	92	10	US-09-764-868-799	Sequence 799, App
23	411	17.3	1130	12	US-10-204-041-4	Sequence 4, Appl1
24	410	17.3	1130	12	US-10-171-889-1	Sequence 1, Appl1
25	410	17.3	1130	12	US-10-263-480-2	Sequence 2, Appl1
26	405	17.1	1567	12	US-10-312-918-2	Sequence 2, Appl1
27	405	17.1	1594	12	US-10-312-918-4	Sequence 4, Appl1
28	392.5	16.5	835	10	US-09-947-199-8	Sequence 8, Appl1
29	385.5	16.2	537	12	US-09-976-782-85	Sequence 85, Appl1
30	385	16.2	1036	10	US-09-771-161A-255	Sequence 255, App
31	385	16.2	1036	10	US-09-771-161A-256	Sequence 256, App
32	385	16.2	1036	12	US-10-318-410-2	Sequence 2, Appl1
33	385	16.2	1040	12	US-10-318-410-5	Sequence 5, Appl1
34	382.5	16.1	251	8	US-08-987-689A-32	Sequence 32, Appl1
35	382.5	16.1	251	15	US-10-292-524-32	Sequence 32, Appl1
36	379.5	16.0	505	9	US-09-977-269-6	Sequence 6, Appl1
37	379.5	16.0	505	10	US-09-982-610-20	Sequence 20, Appl1
38	379.5	16.0	505	10	US-09-977-260-6	Sequence 6, Appl1
39	379.5	16.0	505	11	US-09-977-261-6	Sequence 6, Appl1
40	378.5	15.9	252	12	US-09-976-782-41	Sequence 41, Appl1
41	378.5	15.9	254	12	US-09-976-782-30	Sequence 30, Appl1
42	378.5	15.9	256	12	US-09-863-776-41	Sequence 41, Appl1
43	377	15.9	835	10	US-09-947-199-2	Sequence 2, Appl1
44	376.5	15.9	537	9	US-09-977-269-11	Sequence 11, Appl1
45	376.5	15.9	537	10	US-09-977-260-11	Sequence 11, Appl1

ALIGNMENTS

RESULT 1
US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Action, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver.. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match	100.0%	Score 2375;	DB 9;	Length 455;
Best Local Similarity	100.0%	Pred. No. 2e-175;		
Matches 455;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSLSGASFYQIKFDLDQFENCQGGSPGVSVPRAKWSIQDKEVAVKLLIKIEKAEITLSVL	60	
DB	1	MSLSGASFYQIKFDLDQFENCQGGSPGVSVPRAKWSIQDKEVAVKLLIKIEKAEITLSVL	60	
QY	61	SHRNITIOFYGVILLEPPNYGIVTEYASGLDYDINSNRSEEMDMHIMTWADVAKGMHY	120	
DB	61	SHRNITIOFYGVILLEPPNYGIVTEYASGLDYDINSNRSEEMDMHIMTWADVAKGMHY	120	
QY	121	LHNEAVKVIHDKSRNVIAADGVLTICDGRASFHHTHTMSLVGFFPMVAPEVIO	180	
DB	121	LHNEAVKVIHDKSRNVIAADGVLTICDGRASFHHTHTMSLVGFFPMVAPEVIO	180	
QY	181	LPVSETCDYSGVVMEMTLREVPFKGLEQVAVLVVEKNERLTI PSSCPSPAEILH	240	
DB	181	LPVSETCDYSGVVMEMTLREVPFKGLEQVAVLVVEKNERLTI PSSCPSPAEILH	240	

Db 181 LPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVYKNERLITPSSCPRSPAEHLH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSTLHNKAMRCIEATLERLKLEERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSTLHNKAMRCIEATLERLKLEERD 300
QY 301 LSPFEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTEESNAEMSCOI 360
Db 301 LSPFEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTEESNAEMSCOI 360
QY 361 TATSNGEIGHGNPSSLQAMMLMGFGDIFSNKXGAVVHSGMOINMOAKONSKTTSKRGRK 420
Db 361 TATSNGEIGHGNPSSLQAMMLMGFGDIFSNKXGAVVHSGMOINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 2

US-10-094-749-2477
; Sequence 2477, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUTUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: MASAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2477
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2477

Query Match 100.0%; Score 2375; DB 12; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-175;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGASFOVKRPDDLOFPENCGGSGFSYRAKMTSODKEVAVKKLTKIEKEAEITLSVL 60
Db 1 MSLSGASFOVKRPDDLOFPENCGGSGFSYRAKMTSODKEVAVKKLTKIEKEAEITLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTATDVAKGMY 120
Db 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTATDVAKGMY 120
QY 121 LHMEAPVKVIHRDLKSRNVIIAADGVLTICDPGASRFHNHTTHMSLVGTFPMMAPEVIO 180
Db 121 LHMEAPVKVIHRDLKSRNVIIAADGVLTICDPGASRFHNHTTHMSLVGTFPMMAPEVIO 180

QY 181 LPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVYKNERLITPSSCPRSPAEHLH 240
Db 181 LPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVYKNERLITPSSCPRSPAEHLH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSTLHNKAMRCIEATLERLKLEERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSTLHNKAMRCIEATLERLKLEERD 300
QY 301 LSPFEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTEESNAEMSCOI 360
Db 301 LSPFEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESKTEESNAEMSCOI 360
QY 361 TATSNGEIGHGNPSSLQAMMLMGFGDIFSNKXGAVVHSGMOINMOAKONSKTTSKRGRK 420
Db 361 TATSNGEIGHGNPSSLQAMMLMGFGDIFSNKXGAVVHSGMOINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 3

US-10-106-698-6345
; Sequence 6345, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide;
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6345
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: MISC FEATURE
; NAME/KEY: MISC FEATURE
; LOCATION: (340)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6345

Query Match 59.3%; Score 1409; DB 15; Length 349;
Best Local Similarity 100.0%; Pred. No. 6.7e-101;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGASFOVKRPDDLOFPENCGGSGFSYRAKMTSODKEVAVKKLTKIEKEAEITLSVL 60
Db 57 MSLSGASFOVKRPDDLOFPENCGGSGFSYRAKMTSODKEVAVKKLTKIEKEAEITLSVL 116
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTATDVAKGMY 120
Db 117 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTATDVAKGMY 176
QY 121 LHMEAPVKVIHRDLKSRNVIIAADGVLTICDPGASRFHNHTTHMSLVGTFPMMAPEVIO 180
Db 177 LHMEAPVKVIHRDLKSRNVIIAADGVLTICDPGASRFHNHTTHMSLVGTFPMMAPEVIO 236
QY 181 LPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVYKNERLITPSSCPRSPAEHLH 240
Db 237 LPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVYKNERLITPSSCPRSPAEHLH 296
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSL 269
Db 297 QCWEADAKKRPSFKQIISILESMSNDTSL 325

```

RESULT 4
US-10-354-358-24
/ Sequence 24, Application US/10354358
/ Publication No. US20030157082A1
/ GENERAL INFORMATION:
/ APPLICANT: Millennium Pharmaceuticals, Inc
/ APPLICANT: Hunter, John Joseph
/ APPLICANT: Macbeth, Kyle J.
/ APPLICANT: Tsai, Feng-Ying
/ APPLICANT: Lesoon, Andrea
/ APPLICANT: Lightcap, Eric S.
/ APPLICANT: Williamson, Mark
/ APPLICANT: Rudolph-Owen, Laura A
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
/ TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
/ TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
/ TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
/ TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
/ TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
/ TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
/ FILE REFERENCE: MP102-020P1RNMNMIM
/ CURRENT APPLICATION NUMBER: US/10/354,358
/ PRIOR FILING DATE: 2003-01-30
/ PRIOR APPLICATION NUMBER: US 60/353,600
/ PRIOR FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: US 60/364,517
/ PRIOR FILING DATE: 2002-03-15
/ PRIOR APPLICATION NUMBER: US 60/371,075
/ PRIOR FILING DATE: 2002-04-09
/ PRIOR APPLICATION NUMBER: US 60/371,507
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: US 60/372,984
/ PRIOR FILING DATE: 2002-04-16
/ PRIOR APPLICATION NUMBER: US 60/374,194
/ PRIOR FILING DATE: 2002-04-19
/ PRIOR APPLICATION NUMBER: US 60/382,995
/ PRIOR FILING DATE: 2002-05-24
/ PRIOR APPLICATION NUMBER: US 60/385,023
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: US 60/388,853
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: US 60/389,395
/ PRIOR FILING DATE: 2002-06-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 122
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24
/ LENGTH: 1036
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-354-358-24

Query Match      25.2%; Score 599.5; DB 12; Length 1036;
Best Local Similarity 33.3%; Pred. No. 9.9e-38;
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

```

```

US-10-014-882-2
/ Sequence 2, Application US/10014882
/ Publication No. US20020107384A1
/ GENERAL INFORMATION:
/ APPLICANT: Hu, Yi
/ APPLICANT: Kieke, James
/ APPLICANT: Donoho, Gregory
/ TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding It
/ FILE REFERENCE: LEX-0279-USA
/ CURRENT APPLICATION NUMBER: US/10/014,882
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: US 60/254,744
/ PRIOR FILING DATE: 2000-12-11
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 1036
/ TYPE: PRT
/ ORGANISM: homo sapiens
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(1036)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match      25.2%; Score 599.5; DB 14; Length 1036;
Best Local Similarity 33.3%; Pred. No. 9.9e-38;
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

```

```

QY 325 Q-----SNTPLPLPLAAMSEESYFESKTESNA 354
DB 470 REIDVLERELNLIIFOLNDEKPKYKRGKFKSRKLKDGRIISLPSPQHKITVQASP 529
QY 355 EM-----SCOTTATNSGEHGMNPSLOAMML 380
DB 530 NLDKRSLSNSSSSSPSSPTMMRLRAIQ 559

RESULT 6
US-09-862-027-19
; Sequence 19, Application US/09662027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NO. US20020142428A1e1 kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

Query Match 25.0%; Score 593; DB 10; Length 394;
Best Local Similarity 39.3%; Pred. No. 9.3e-38;
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

QY 15 DLQFPNCGGSGFGSYRAKMTISQDKRVAVK-----KLKIEKEAEILSVLSH 62
DB 2 ELTLEELIIGGFGKYRAFWIGD--EVAVKARHPDEDISQTIENVQGEAKLFAMLKX 59
QY 63 RNIOGYGVILEPPNYGIYEVASLSGLYDINSNSEMDMHMTATDVAKGMHYLH 122
DB 60 PNIALRGVCLKEPNCLVMEFARGGPLNRVLSGKR---IPPIILVMAVQIARGMNYLH 116
QY 123 MEAPVVIHRDLKSRNVV-----AADGVLCIDFGASRPHNHTTHSLVGTFFPMA 174
DB 117 DEIVLPIIHRDLSSNILLIQKVENGDLSNKILKITDFGLARWHRTTKMSAGTYANMA 176
QY 175 PEVIOGLPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVKNERLTIIPSSCPBS 234
DB 177 PEVIRASMSKSGSDVMSYGVLLWELLTGEVPRGIDGLRVAAGVANMKLALPIPTCPEP 236
QY 235 FAELHQCWEADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHAKAEKCEIATLERL 294
DB 237 FAELMEDCWNPDHSPSPFTNILDQTLTIEESGFEMPKDSFRLQDNMKGHEIQEMFDOL 296
QY 295 KLERDLSFKEOEL-----KERERRLKMEOKLTEQ 325
DB 297 RAKEKELRTWEELTRALQOKQOEELRLRREOELAE 334

RESULT 7
US-10-288-798-12
; Sequence 12, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YOB, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;

```

```

; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dzung Anna M.; LAB, Preeti G.;
; APPLICANT: RAMKOMAR, Jayalakshi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 55052990CD1
US-10-288-798-12

Query Match 24.8%; Score 590; DB 12; Length 1097;
Best Local Similarity 39.0%; Pred. No. 5.8e-37;
Matches 134; Conservative 60; Mismatches 118; Indels 32; Gaps 5;

QY 9 VOIKEDDLQFPNCGGSGFGSYRAKMTISQDKRVAVK-----KLKIEKEAEI 56
DB 137 LEIDFAELLEELIIGGFGKYRAFWIGD--EVAVKARHPDEDISQTIENVQGEAKL 194
QY 57 LSVLSHRNIOGYGVILEPPNYGIYEVASLSGLYDINSNSEMDMHMTATDVAK 116
DB 195 FAMLKPNIALRGVCLKEPNCLVMEFARGGPLNRVLSGKR---IPPIILVMAVQIAR 251
QY 117 GMHYLMEAPVVIHRDLKSRNVV-----AADGVLCIDFGASRPHNHTTHSLVG 168
DB 252 GMNYLDEIVLPIIHRDLSSNILLIQKVENGDLSNKILKITDFGLARWHRTTKMSAG 311
QY 169 TEPWMAPEVIOGLPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAMLVKNERLTI 228
DB 312 TYANMAPEVIRASMSKSGSDVMSYGVLLWELLTGEVPRGIDGLRVAAGVANMKLALPI 371
QY 229 SSCPSFAELHQCWEADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHAKAEKCEI 288
DB 372 STCPPEFAELMEDCWNPDHSPSPFTNILDQTLTIEESGFEMPKDSFRLQDNMKGHEI 431
QY 289 ATLRLKLERDLSFKEOEL-----KERERRLKMEOKLTEQ 325
DB 432 EMPDQRAKEKELRTWEELTRALQOKQOEELRLRREOELAE 475

RESULT 8
US-10-143-133-2
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use

```



```

Db      209  FKGVCQTQAPCYCILMEFCAQGQYEVLRAGRVPVPSL--LVDMWGIAGGNNYLHLH---263
Qy      128  KVIHRDLKSRNNVIAADGVLCIDCFGASR-FHNHTTHMSLVGTFPMWAPREVISOIQLPVSET 186
Db      264  KIIHRDLKSPNMLITVDVVKISDFGTSKELSDKSTMSFAGTVAMAPREVINEPVSEK 323
Qy      187  CDTYSYGVVLMEMLTREVFPKGLBGLQVAMLVVEKNERLTTPSSCPSPFAELLHQCEAD 246
Db      324  VDIWISFGVVLMEILTGEIPYKDVDSAIIWGVSNSLHLPVPSGCPDGFKILLRQCWNSK 383
Qy      247  AKRPSFKQIISILBMSNDT-SLPDKNSFLHNKAEWCEIATLERLK-----KLE 298
Db      384  PRNRPSPRQILHLHDIASADVLPSTPGE--TYFKSQAEWREEVKLFHEKIKSGTCLHRL 441
Qy      299  RDISFKEQE-----LKERERRLKMEOKLTEOSNTPLLLPLARMBEESYFESKTEE 350
Db      442  BELWMRRREBLRHADIREH-----YERKLEFRANN--LYMELNALMLQ---LELKERE 489

```

RESULT 11

```

US-10-153-668-324
; Sequence 324, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MURAMATSU, Akio
; APPLICANT: ISHIZAMA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153, 668
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
; LENGTH: 859
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-153-668-324

```

```

Query Match      23.8%; Score 565.5; DB 15; Length 859;
Best Local Similarity 37.8%; Pred. No. 3,3e-35;
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

Qy      10  QIKFDDLOFFENGGGSGSVYRAKMTISQDEKVAVKKLIKIEKAEI--LSVLSHRNIIQ 67
Db      119  EVPFEEIIDLQWVGSGAQAVFLGRF--HGEVAVKVRDL-KETDIKHLRKIKHPNIIT 175
Qy      68  FYGVILEPPNYGIVTEYASLGSLYDYNRSREEMDMHIMTAVDVAKGMYLHMEAPV 127
Db      176  FKGVCQTQAPCYCILMEFCAQGQYEVLRAGRVPVPSL--LVDMWGIAGGNNYLHLH---230
Qy      128  KVIHRDLKSRNNVIAADGVLCIDCFGASR-FHNHTTHMSLVGTFPMWAPREVISOIQLPVSET 186
Db      231  KIIHRDLKSPNMLITVDVVKISDFGTSKELSDKSTMSFAGTVAMAPREVINEPVSEK 290
Qy      187  CDTYSYGVVLMEMLTREVFPKGLBGLQVAMLVVEKNERLTTPSSCPSPFAELLHQCEAD 246
Db      291  VDIWISFGVVLMEILTGEIPYKDVDSAIIWGVSNSLHLPVPSGCPDGFKILLRQCWNSK 350
Qy      247  AKRPSFKQIISILBMSNDT-SLPDKNSFLHNKAEWCEIATLERLK-----KLE 298

```

```

Db      351  PRNRPSPRQILHLHDIASADVLPSTPGE--TYFKSQAEWREEVKLFHEKIKSGTCLHRL 408
Qy      299  RDISFKEQE-----LKERERRLKMEOKLTEOSNTPLLLPLARMBEESYFESKTEE 350
Db      409  BELWMRRREBLRHADIREH-----YERKLEFRANN--LYMELNALMLQ---LELKERE 456

```

RESULT 12

```

US-10-153-668-438
; Sequence 438, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MURAMATSU, Akio
; APPLICANT: ISHIZAMA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153, 668
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 438
; LENGTH: 892
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-153-668-438

```

```

Query Match      23.8%; Score 565.5; DB 15; Length 892;
Best Local Similarity 37.8%; Pred. No. 3,5e-35;
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

Qy      10  QIKFDDLOFFENGGGSGSVYRAKMTISQDEKVAVKKLIKIEKAEI--LSVLSHRNIIQ 67
Db      152  EVPFEEIIDLQWVGSGAQAVFLGRF--HGEVAVKVRDL-KETDIKHLRKIKHPNIIT 208
Qy      68  FYGVILEPPNYGIVTEYASLGSLYDYNRSREEMDMHIMTAVDVAKGMYLHMEAPV 127
Db      209  FKGVCQTQAPCYCILMEFCAQGQYEVLRAGRVPVPSL--LVDMWGIAGGNNYLHLH---263
Qy      128  KVIHRDLKSRNNVIAADGVLCIDCFGASR-FHNHTTHMSLVGTFPMWAPREVISOIQLPVSET 186
Db      264  KIIHRDLKSPNMLITVDVVKISDFGTSKELSDKSTMSFAGTVAMAPREVINEPVSEK 323
Qy      187  CDTYSYGVVLMEMLTREVFPKGLBGLQVAMLVVEKNERLTTPSSCPSPFAELLHQCEAD 246
Db      324  VDIWISFGVVLMEILTGEIPYKDVDSAIIWGVSNSLHLPVPSGCPDGFKILLRQCWNSK 383
Qy      247  AKRPSFKQIISILBMSNDT-SLPDKNSFLHNKAEWCEIATLERLK-----KLE 298
Db      384  PRNRPSPRQILHLHDIASADVLPSTPGE--TYFKSQAEWREEVKLFHEKIKSGTCLHRL 441
Qy      299  RDISFKEQE-----LKERERRLKMEOKLTEOSNTPLLLPLARMBEESYFESKTEE 350
Db      442  BELWMRRREBLRHADIREH-----YERKLEFRANN--LYMELNALMLQ---LELKERE 489

```

RESULT 13

```

US-09-862-027-18
; Sequence 18, Application US/09862027

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:17:01 ; Search time 19.4444 Seconds
(without alignments)
990.074 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375

Sequence: 1 MSLSASFVQIKFDDIQFFE.....GDDDDDDGDEERDNDMNSE 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2375	100.0	455	3	US-09-221-235-5	Sequence 5, Appl1
2	2375	100.0	455	3	US-09-221-928-5	Sequence 5, Appl1
3	2375	100.0	455	3	US-09-221-527-5	Sequence 5, Appl1
4	2375	100.0	455	3	US-09-221-236-5	Sequence 5, Appl1
5	2375	100.0	455	3	US-09-221-416-5	Sequence 5, Appl1
6	2375	100.0	455	3	US-09-221-245-5	Sequence 5, Appl1
7	2375	100.0	455	3	US-09-163-115-5	Sequence 5, Appl1
8	2375	100.0	455	3	US-09-221-528-5	Sequence 5, Appl1
9	2375	100.0	455	3	US-09-593-553-5	Sequence 5, Appl1
10	2375	100.0	455	3	US-09-221-237-5	Sequence 5, Appl1
11	2375	100.0	455	4	US-09-399-588-2	Sequence 2, Appl1
12	593	25.0	394	4	US-09-345-473E-19	Sequence 19, Appl1
13	565.5	23.8	859	1	US-08-395-580-2	Sequence 2, Appl1
14	565.5	23.8	859	5	PCT-US95-02799-2	Sequence 2, Appl1
15	565.5	23.8	328	4	US-09-345-473E-18	Sequence 18, Appl1
16	545.5	23.0	821	1	US-07-928-464-2	Sequence 2, Appl1
17	488	20.5	821	1	US-08-003-311B-2	Sequence 2, Appl1
18	488	20.5	821	1	US-08-261-432-2	Sequence 2, Appl1
19	488	20.5	821	5	PCT-US93-07347-2	Sequence 2, Appl1
20	487	20.5	263	3	US-09-035-706-5	Sequence 5, Appl1
21	487	20.5	263	3	US-08-955-841-5	Sequence 5, Appl1
22	487	20.5	263	4	US-09-390-425-5	Sequence 5, Appl1
23	487	20.5	263	4	US-09-566-906-5	Sequence 5, Appl1
24	481	20.3	579	4	US-09-529-279-4	Sequence 4, Appl1
25	481	20.3	579	4	US-10-158-895-4	Sequence 4, Appl1
26	481	20.3	590	4	US-09-529-279-15	Sequence 15, Appl1
27	481	20.3	590	4	US-09-529-279-15	Sequence 15, Appl1

28	481	20.3	590	4	US-10-158-895-15	Sequence 15, Appl1
29	401.5	16.9	1584	2	US-09-457-040B-27	Sequence 27, Appl1
30	397.5	16.7	275	2	US-08-701-191A-36	Sequence 36, Appl1
31	393.5	16.6	261	2	US-07-857-224B-59	Sequence 59, Appl1
32	392.5	16.5	835	4	US-09-458-457-8	Sequence 8, Appl1
33	391.5	16.5	261	2	US-07-857-224B-60	Sequence 60, Appl1
34	379.5	16.0	505	1	US-08-222-616-20	Sequence 20, Appl1
35	379.5	16.0	505	4	US-08-446-648-20	Sequence 20, Appl1
36	379.5	16.0	505	4	US-08-492-723-2	Sequence 2, Appl1
37	379.5	16.0	505	5	PCT-US95-04228-20	Sequence 20, Appl1
38	377.5	15.9	506	4	US-08-426-509A-6	Sequence 6, Appl1
39	377.5	15.9	511	4	US-08-232-545-6	Sequence 6, Appl1
40	377.5	15.9	511	5	PCT-US95-05008-6	Sequence 6, Appl1
41	377	15.9	835	3	US-09-291-839-2	Sequence 2, Appl1
42	377	15.9	835	4	US-09-458-457-2	Sequence 2, Appl1
43	376.5	15.9	537	4	US-08-426-509A-11	Sequence 11, Appl1
44	376.5	15.9	537	4	US-08-232-545-11	Sequence 11, Appl1
45	376.5	15.9	537	5	PCT-US95-05008-11	Sequence 11, Appl1

ALIGNMENTS

RESULT 1									
US-09-221-235-5									
; Sequence 5, Application US/09221235									
; Patent No. 6043040									
; GENERAL INFORMATION:									
; APPLICANT: Acton, Susan									
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR									
; FILE REFERENCE: NMI-050									
; CURRENT APPLICATION NUMBER: US/09/221,235									
; CURRENT FILING DATE: 1998-12-28									
; EARLIER APPLICATION NUMBER: 09/163,115									
; EARLIER FILING DATE:									
; NUMBER OF SEQ ID NOS: 15									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 5									
; LENGTH: 455									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; US-09-221-235-5									
Query Match									
Best Local Similarity 100.0%; Score 2375; DB 3; Length 455;									
Matches 455; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MSLSASFVQIKFDDIQFFENC	GGGFGSVYAKMTSDQKEYAVKKLTIKREAITLVY	60					
DB	1	MSLSASFVQIKFDDIQFFENC	GGGFGSVYAKMTSDQKEYAVKKLTIKREAITLVY	60					
QY	61	SHRNIIQFVGLLEPPENYIV	TEYASIGSLVYINSNRSEEDMDHIMTATDVAKGMHY	120					
DB	61	SHRNIIQFVGLLEPPENYIV	TEYASIGSLVYINSNRSEEDMDHIMTATDVAKGMHY	120					
QY	121	LHNEAPVKIHRDLSRNVII	ADGVLYKICDFGASFFNHHTHMSLVGFPPMAAPVLOS	180					
DB	121	LHNEAPVKIHRDLSRNVII	ADGVLYKICDFGASFFNHHTHMSLVGFPPMAAPVLOS	180					
QY	181	LPVSEICDYISGVVLMWLT	REVPKGLGEGVQVAMLVYKNERLTISSCPSPFELH	240					
DB	181	LPVSEICDYISGVVLMWLT	REVPKGLGEGVQVAMLVYKNERLTISSCPSPFELH	240					
QY	241	QCEWADAKRPSFKQIISIL	ESMSNDTSLPDKNSFLHNKAEMRCEIEATLERKLEBD	300					
DB	241	QCEWADAKRPSFKQIISIL	ESMSNDTSLPDKNSFLHNKAEMRCEIEATLERKLEBD	300					
QY	301	LSFKEOELKERERRLKME	QOKLTBQSNTPLLPLAARMEESYFBSKTESNAEMSCOI	360					
DB	301	LSFKEOELKERERRLKME	QOKLTBQSNTPLLPLAARMEESYFBSKTESNAEMSCOI	360					
QY	361	TTSNCEGGMPSLOAMLMG	FQDFFSMNKGAVVHSGMQUINMOAKONSSKTTSRGRK	420					
DB	361	TTSNCEGGMPSLOAMLMG	FQDFFSMNKGAVVHSGMQUINMOAKONSSKTTSRGRK	420					

Db 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNMKAGAVHSGMOMQAKONSSKTTSKRGK 420
QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 2
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Accion, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFGSVYRAKWSQDEKVAVKKLKIEKAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGFGSVYRAKWSQDEKVAVKKLKIEKAEILSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDIYNSRSEEMDMHIMTATDVAKGMHY 120
Db 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDIYNSRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMMAPEVIO 180
Db 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMMAPEVIO 180
QY 181 LPVSETCDTYSYGVVLMEMLTREVPPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
Db 181 LPVSETCDTYSYGVVLMEMLTREVPPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
QY 241 QCWEADAKKRPSPKQIISILESMSNDTSLPDKCNSTFLHKAEMRCIEATLERLKLKLERD 300
Db 241 QCWEADAKKRPSPKQIISILESMSNDTSLPDKCNSTFLHKAEMRCIEATLERLKLKLERD 300
QY 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNAEMSCQI 360
Db 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNAEMSCQI 360
QY 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNMKAGAVHSGMOMQAKONSSKTTSKRGK 420
Db 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNMKAGAVHSGMOMQAKONSSKTTSKRGK 420
QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 3
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Accion, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFGSVYRAKWSQDEKVAVKKLKIEKAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGFGSVYRAKWSQDEKVAVKKLKIEKAEILSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDIYNSRSEEMDMHIMTATDVAKGMHY 120
Db 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDIYNSRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMMAPEVIO 180
Db 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMMAPEVIO 180
QY 181 LPVSETCDTYSYGVVLMEMLTREVPPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
Db 181 LPVSETCDTYSYGVVLMEMLTREVPPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
QY 241 QCWEADAKKRPSPKQIISILESMSNDTSLPDKCNSTFLHKAEMRCIEATLERLKLKLERD 300
Db 241 QCWEADAKKRPSPKQIISILESMSNDTSLPDKCNSTFLHKAEMRCIEATLERLKLKLERD 300
QY 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNAEMSCQI 360
Db 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNAEMSCQI 360
QY 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNMKAGAVHSGMOMQAKONSSKTTSKRGK 420
Db 361 TATSNBEGHGMNPSLOAMLMGFGDIFSNMKAGAVHSGMOMQAKONSSKTTSKRGK 420
QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 4
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Accion, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFGSVYRAKWSQDEKVAVKKLKIEKAEILSVL 60

```

Db      1  |||||
        1  MSSLGASFYQIKRDDLOPFENCGGSGFSGYVRAKWTISQDKEVAVKLLKTEKAEITLSVL 60
Qy      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Db      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Qy      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Db      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Qy      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Db      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Qy      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300
Db      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300
Qy      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Db      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Qy      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Db      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Qy      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 5
US-09-221-416-5
Sequence 5, Application US/09221416
Patent No. 6153417
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,416
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRN
ORGANISM: Homo sapiens
US-09-221-416-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1  |||||
        1  MSSLGASFYQIKRDDLOPFENCGGSGFSGYVRAKWTISQDKEVAVKLLKTEKAEITLSVL 60
Db      1  |||||
        1  MSSLGASFYQIKRDDLOPFENCGGSGFSGYVRAKWTISQDKEVAVKLLKTEKAEITLSVL 60
Qy      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Db      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Qy      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Db      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Qy      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Db      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Qy      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300

```

```

Db      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300
Qy      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Db      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Qy      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Db      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Qy      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 6
US-09-221-245-5
Sequence 5, Application US/09221245
Patent No. 6180358
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,245
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: US 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRN
ORGANISM: Homo sapiens
US-09-221-245-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1  |||||
        1  MSSLGASFYQIKRDDLOPFENCGGSGFSGYVRAKWTISQDKEVAVKLLKTEKAEITLSVL 60
Db      1  |||||
        1  MSSLGASFYQIKRDDLOPFENCGGSGFSGYVRAKWTISQDKEVAVKLLKTEKAEITLSVL 60
Qy      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Db      61  |||||
        61  SHRNIIQFYGVIIPEPNYGI VTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
Qy      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Db      121  |||||
        121  LHMEAPVKIHRDLKSRNVIIADGVKICDPGASRFHNHTTMSLVGTFFPMMAPEVIQS 180
Qy      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Db      181  |||||
        181  LPVSETCDTYSYGVVLMEMLTREVPEFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELI 240
Qy      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300
Db      241  |||||
        241  QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHAKAEWRCIEATLERLKLEERD 300
Qy      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Db      301  |||||
        301  LSPKEQLERERRRLKMEQKLTQOSNTPLPLAARMSESYFESKTESNSAEMSCOI 360
Qy      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Db      361  |||||
        361  TATSNCEGHGMNPSIQAMMLMGFDIFSNMKAGAVHSGMOIMQAKONSKTTSKRGRK 420
Qy      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421  |||||
        421  KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 7

US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
DB 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
QY 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180
DB 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180
QY 181 LPVSETCDITYSGVVLWEMLTREVPFKGLEQLVAMLVVEKNERLTI PSSCPRSPFAELH 240
DB 181 LPVSETCDITYSGVVLWEMLTREVPFKGLEQLVAMLVVEKNERLTI PSSCPRSPFAELH 240
QY 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCN SFLHKAEMRCEIATLERLKLERD 300
DB 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCN SFLHKAEMRCEIATLERLKLERD 300
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLP LAARMEESYFESKTESNSAEMSCQI 360
DB 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLP LAARMEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFQDIFSNNKAGAVMHSQMOMQAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFQDIFSNNKAGAVMHSQMOMQAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455

RESULT 8

US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens

US-09-221-528-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
DB 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
QY 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180
DB 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180
QY 181 LPVSETCDITYSGVVLWEMLTREVPFKGLEQLVAMLVVEKNERLTI PSSCPRSPFAELH 240
DB 181 LPVSETCDITYSGVVLWEMLTREVPFKGLEQLVAMLVVEKNERLTI PSSCPRSPFAELH 240
QY 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCN SFLHKAEMRCEIATLERLKLERD 300
DB 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCN SFLHKAEMRCEIATLERLKLERD 300
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLP LAARMEESYFESKTESNSAEMSCQI 360
DB 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLP LAARMEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFQDIFSNNKAGAVMHSQMOMQAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFQDIFSNNKAGAVMHSQMOMQAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455

RESULT 9

US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match 100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
DB 1 MSSLGASFWQIKFDDLOFPENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAETLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATDVAKGMHY 120
QY 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180
DB 121 LHMEAPKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMMAPEVIO 180

```

Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNHTTMSLVGTFPMMAPEVIO 180
Qy      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Db      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Qy      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Db      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Qy      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Db      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Qy      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420
Db      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420
Qy      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 10

```

US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-26
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

```

```

Query Match      100.0%; Score 2375; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MSSLGASFWQIKRDDLOFFENCAGGSGFSYVYRAKWTISQDEKAVVKKLTIKEKAEIISVL 60
Db      1 MSSLGASFWQIKRDDLOFFENCAGGSGFSYVYRAKWTISQDEKAVVKKLTIKEKAEIISVL 60
Qy      61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYDINSNRSEMDMIMTAVTAVAKGMHY 120
Db      61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYDINSNRSEMDMIMTAVTAVAKGMHY 120
Qy      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNHTTMSLVGTFPMMAPEVIO 180
Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNHTTMSLVGTFPMMAPEVIO 180
Qy      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Db      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Qy      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Db      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Qy      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Db      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Qy      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420
Db      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420

```

```

Db      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420

```

```

Qy      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 11

```

US-09-399-588-2
; Sequence 2, Application US/09399588
; Patent No. 6511825
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Rosamaria
; APPLICANT: Callow, Marinella
; TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids
; FILE REFERENCE: 1044-US
; CURRENT APPLICATION NUMBER: US/09/399,588
; CURRENT FILING DATE: 1999-09-20
; EARLIER APPLICATION NUMBER: 60/104,088
; EARLIER FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Human J42
US-09-399-588-2

```

```

Query Match      100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.9e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MSSLGASFWQIKRDDLOFFENCAGGSGFSYVYRAKWTISQDEKAVVKKLTIKEKAEIISVL 60
Db      1 MSSLGASFWQIKRDDLOFFENCAGGSGFSYVYRAKWTISQDEKAVVKKLTIKEKAEIISVL 60
Qy      61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYDINSNRSEMDMIMTAVTAVAKGMHY 120
Db      61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYDINSNRSEMDMIMTAVTAVAKGMHY 120
Qy      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNHTTMSLVGTFPMMAPEVIO 180
Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNHTTMSLVGTFPMMAPEVIO 180
Qy      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Db      181 LPVSEICDTYISYGVVWEMLTREVPRKGLGLOVAVLVEKNERLTIIPSCPRSPFALLH 240
Qy      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Db      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNFSLHNAEWRCEIATLERLTKLERD 300
Qy      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Db      301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTESNGAEMSCQI 360
Qy      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420
Db      361 TATSNEGHGMNPISLOAMLMGFGDIFSNMKAGAVHSGQIMQAKONSKTTSKRGRK 420
Qy      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455
Db      421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```

RESULT 12

```

US-09-345-473E-19
; Sequence 19, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof

```

```

; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-345-473E-19

Query Match      25.0%; Score 593; DB 4; Length 394;
Best Local Similarity 39.3%; Pred. No. 4,6e-42;
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

QY 15 DLOFFNCGGSGSYRAKWIQDKEVAVK-----KLKIEKAEILSVLSH 62
DB 2 ELTLEETIGIGFGKVFARFWDG--EVAVKARHPDEDISQTIENVROEAKLFAMLK 59
QY 63 RNIOFGVILEPNYGIVTEVASLGLDYINSRSEMDMDHMTWATDVAKGMHYLH 122
DB 60 PNIIALRGVCLKEPNLCLVNEFARGGFLNRLVSGKR--IPDILVMWAVQIARGMYLH 116
QY 123 MEAPVYIHRDLKSRNVI-----AADGLKICDFGASRPHNHTTHMSLVGTFPWMA 174
DB 117 DEAIVIPPIHRDLKSNLILQKVENGLSNKILKITDFGLAREMHRRTTKSAAGTYAMA 176
QY 175 PEVIQSLPVSETCDTYSYGVLMEMLTREVPFKGLEGLQVAMLVENKERTLPSSCPRS 234
DB 177 PEVIARSMFSGKSDVMSYGVLMELTLEGEVFPFGIDGLRVAQVAMKLLALPIPTCPEP 236
QY 235 FAEILLQGWADAKKRPSEFKQIISLESMSNDTSLPDKNSFLHNKAEMCEIATLERL 294
DB 237 FAKLMEDKMPDPHSRPSFTNILDQTLTIESGFEMPKDSFHCLOQNMWHEIQEMFDQL 296
QY 295 KKLERDLSFKEOEL-----KERERRLKMMEQKLTQ 325
DB 297 RAKEKELRTWEELTRALQCKQOEELRRBOELMR 334

RESULT 13
US-08-205-018-2
; Sequence 2, Application US/08205018
; Patent No. 5554523
; GENERAL INFORMATION:
; APPLICANT: Reddy, Usharani R.
; TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; ADDRESSEE: No. 5554523is
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,018
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaumont, Rebecca R.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-205-018-2

Query Match      23.8%; Score 565.5; DB 1; Length 668;
Best Local Similarity 37.8%; Pred. No. 1.9e-39;
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKPDLQRFNCGGSGPSYRAKWIQDKEVAVYKLLIKKEKEL--LSVLSHNIIQ 67
DB 119 EVFPEEILDLQWVGSQAQAVFLGR--HGEVAVKRVKDL--KETDIKRLRKHPNIT 175
QY 68 FGVILEPNYGIVTEVASLGLDYINSRSEMDMDHMTWATDVAKGMHYLMEAPV 127
DB 176 FKGVCTQACYCTILMEFCQGLYELVRAGRVPYPSL--LVDSMGIAGMYLHILH--- 230
QY 128 KYIHRDLKSRNVIADGVLYKICDFGASR--FHNHTTHMSLVGTFPWMADEVYQSLPVSET 186
DB 231 KTIHRDLKSPNMLITYDVVKISDFGTSKELSDKSTKMSFAGTVAMAEVIRNEPVSER 290
QY 187 CDTYSYGVLMEMLTREVPFKGLEGLQVAMLVENKERTLPSSCPRSFAELLHQWMEAD 246
DB 291 VIVMSGVVLMELTGEIRYKQVDSALIWGNSLHPVSSCPDGKILLRQCMNSK 350
QY 247 AKKRPSEFKQIISLESMSNDT--SLPDKNSFLHNKAEMCEIATLERLK-----KLE 298
DB 351 PKRPSFRILHLHDIASADVLPQEE--TYRKSQEMREEVKLMHEKIKSGTCLHRL 408
QY 299 RDLSEFKOE-----LKERERRLKMMEQKLTQESNTPLLPLAARSEESYFESKTEE 350
DB 409 BELVMRRREELRHADIREH---YERKLERANN--LYVELNALMLQ---LEIKERE 456

RESULT 14
US-08-395-580-2
; Sequence 2, Application US/08395580
; Patent No. 5676945
; GENERAL INFORMATION:
; APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
; APPLICANT: Hospital of Philadelphia
; TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945is
; ADDRESSEE: No. 5676945
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 Kb
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,580
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/205,018
; FILING DATE: 01-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca L. Ralph (Formerly Gaumont)
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:16:46 ; Search time 11.2593 Seconds
(without alignments) 2109.700 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277

Sequence: 1 YRAKMTSQDKYAVVKKLLKTI.....SILEMSNDTSLPDKCNSEL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	511.5	39.3	406	2	T52626
2	510.5	39.3	394	2	TJ00229
3	495.5	38.1	328	2	T16747
4	490.5	37.7	954	1	S68178
5	490	37.7	668	2	JC2363
6	480	37.7	888	2	A55318
7	488	37.5	888	2	JC5399
8	482.5	37.1	847	1	A53800
9	480	36.9	886	2	T48544
10	475.5	36.6	1030	2	P96763
11	460.5	35.4	462	2	S29851
12	457.5	35.2	380	2	T01451
13	453	34.8	982	2	T06576
14	448	34.5	412	2	T10671
15	442	34.0	821	2	T48400
16	436.5	33.6	1015	2	T00726
17	434	33.4	848	2	B87950
18	434	33.4	855	2	T20082
19	431.5	33.2	553	2	T04683
20	431	33.2	829	2	T07406
21	427	32.8	546	2	T12955
22	420.5	32.3	546	2	D84555
23	418	32.2	407	2	G84635
24	417	32.1	736	2	T05137
25	403.5	31.0	968	2	F86316
26	402	30.9	963	2	T09911
27	396.5	30.5	545	2	T05675
28	395	30.4	567	2	JC5957
29	395	30.4	579	2	JC5955

30	395	30.4	606	2	JC5956	transforming growt
31	381.5	29.3	738	2	F96701	hypothetical prote
32	369.5	28.4	1130	1	TVHUA	protein-tyrosine k
33	369	28.4	1257	2	T00486	serine/threonine-s
34	368.5	28.3	981	1	POWVGM	gag-ab1 polypeptel
35	368.5	28.3	1123	2	A39962	kinase-related tra
36	367	28.2	364	2	G71410	probable protein k
37	365.5	28.1	1146	2	B35962	protein-tyrosine k
38	365.5	28.1	1182	2	A35962	protein-tyrosine k
39	364.5	28.0	1520	1	TVFFA	protein-tyrosine k
40	364.5	27.9	357	2	C84856	probable protein k
41	362.5	27.9	391	2	T48115	protein kinase ATM
42	362	27.8	842	2	T32258	hypothetical prote
43	358.5	27.6	1147	2	F86297	hypothetical prote
44	355.5	27.3	1584	2	T18276	protein-tyrosine k
45	354.5	27.3	1171	2	T12956	hypothetical prote

ALIGNMENTS

RESULT 1

T52626

probable mitogen-activated protein kinase MAPK delta-1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #ext_change 08-Dec-2000

C:Accession: T52626

R:Ouanic, S.; Hamel, A.; Lepince, A.S.; Tregear, J.W.; Krejs, M.; Henry, Y.

Gene 229, 171-81, 1999

A:Title: Characterisation of novel plant genes encoding MEKK/STK11 and RAF-related protei

A:Reference number: Z24447; MUID:99196996; PMID:10095117

A:Accession: T52626

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-406 <JOU>

A:Molecule type: mRNA

A:Cross-references: EMBL:Y14199; NID:G2253009; PIDN:CAA74591.1; PID:G2253010

A:Experimental source: Cultivar Columbia

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 39.3%; Score 511.5; DB 2; Length 406;
Best Local Similarity 40.5%; Pred. No. 1.3e-23;
Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY	1	YRAKMTSQDKYAVVKKLL-----KIEKEAEILSVLSHRNIIQFGVLEPPNYGI	50
DB	150	YAEW--NGTEVAVKKFLDQDFSGDALTFKSEIIMRLRPNVVLFWGAVTRPPNFSI	207
QY	51	VTEVYSLGSLYDYINSNRSEEMDMHIMTAVDAKGMHYLHMEAPVKVIHRDLKSRNV	110
DB	208	LIEFLPRGSLYRLH-RPHQIDDKRMALDVAKGMVYLTSHPT-VVHDLKSPNLL	265
QY	111	IAADGVLLKIDFGASRFNHT--THNSLVGTFPMMAPEVIOGLPVSETCVTSYGVVLM	168
DB	266	VKNWVAVKVCDFGLSMKHTYLSKSTAGTEEMWAPVLRBNPANEKCDVVSFGVILM	325
QY	169	MUTREVPFEGSLQYAMVVEKNERLTTPSSCRPSFALLIQCMADAKKPPSFQIIS	228
DB	326	LATSRVPMKGLPMQVAVGFQNRRLTIPDDIDLTVAQIIRECQTEHRLRPSFTQIMQ	385
QY	229	ILESMSNDTSLPDKCNS 245	
DB	386	SLKRLQG-LNISNRANT 401	

RESULT 2

JU0229

mixed-kinase protein kinase 1 - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #ext_change 17-Mar-1999

C:Accession: S33467; JU0229

R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.

Eur. J. Biochem. 213, 701-710, 1993

A:Title: Identification of a new family of human epithelial protein kinases containing ty

A:Reference number: S32467; MUID:93238756; PMID:8477742
 A:Accession: S32467
 A:Molecule type: mRNA
 A:Residues: 1-394 <DO>
 C:Genetics:
 A:Gene: GDB:MLK1
 A:Cross-references: GDB:141921; OMIM:600136
 A:Map position: 14q24.3-14q31
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein kinase
 F:1-269/Domain: protein kinase homology <KIN>
 F:1-268/Domain: catalytic <CAT>
 F:9-17/Region: protein kinase ATP-binding motif
 F:289-310/Region: leucine zipper motif
 F:324-345/Region: leucine zipper motif
 F:354-368/Region: basic

Query Match 39.3%; Score 510.5; DB 2; Length 394;
 Best Local Similarity 42.7%; Pred. No. 1.5e-23;
 Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAKWSQDEKVAVK-----KLKIKKAEILSVLSHRNIQFYGVILEPPNY 48
 DB 18 YRAFWICD--EVAVKARHPDEDISQTIENVRQEAFLFAMLKHPNIIALRGVCLKEPNL 75
 QY 49 GYTEVASIGSLVDYINSNSEEMDMHMTWATDVAKGMYLHMEAPYKVIHRDLKSRN 108
 DB 76 CLWEPFARGGPLNRVLSGKR---IPPIIVMVAQVLRGNVYLHDELYVPIHRDLKSSN 132
 QY 109 VVI-----AADGVLIKIDFGASRPHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTY 160
 DB 133 ILLQKVENDDLSNKLIKIDFGARHMTTMSAAGTAMAPVIRASMSKGSDDW 192
 QY 161 SYGVLMEMITREVPFKGLGLOYAMLVKNERLTIPSSCPSPFELHQCHEADAKR 220
 DB 193 SYGVLMELITGEVFPFGIGDLRVAYGVANMKLALPIPTCPPEPAKLMDCNNPDPHSR 252
 QY 221 PSFKQIISLESM 233
 DB 253 PSFTNIDQLTTI 265

RESULT 3

TI6747
 hypothetical protein R13F6.7 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
 C/Accession: T16747
 R/Miller, N.
 submitted to the EMBL Data Library, April 1994
 A:Description: The sequence of C. elegans cosmid R13F6.
 A:Reference number: Z18570
 A:Accession: T16747
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-328 <MTL>
 A:Cross-references: EMBL:U00046; NID:g470358; PID:g470364; PIDN:AAC47047.1; GSPDB:GN0002
 A:Experimental source: strain Bristol N2; clone R13F6
 C:Genetics:
 A:Gene: CESP:R13F6.7
 A:Map position: 3
 A:Introns: 20/3; 160/3; 222/2; 286/2
 C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology

Query Match 38.1%; Score 495.5; DB 2; Length 328;
 Best Local Similarity 41.0%; Pred. No. 9.7e-23;
 Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

QY 1 YRAKWSQD--KEVAVKLIKIEKEKELISVLSHRNIQFYGVILEPPN-YGVITFYAS 56
 DB 67 FSGNWTLPDQSQRITALKVFVEKEKELISKRHNKIQFYGICRATGDFVITVEYAE 126
 QY 57 LGSIVDYINSRSEEM-----DMDHMTWATDVAKGMYLHMEAPYKVIHRDLKSRN 109

DB 127 KSLVDYFIHSESGCFASSSGNSFDVYVXMASQIASGIQVLYHVAVDPIIHRDLKSKV 186
 QY 110 VTADGVLIKIDFGASRPHNHT-TMSLVGTFPMMAPE-VIOSLVSFECTDYISGVLYM 167
 DB 187 VLDKNLVCKICDFGTSKDLTHSCTAPSWGTVAMMSPEIILQSEGLTATDWSYGVLYM 246
 QY 168 EMLTREVPFKGLEGVAMLVYKNERLTIPSSCPSPFELHQCHEADAKRPSFKQII 227
 DB 247 ELTSKEVPKDYSEFRIFMTIQSGITLAIPTSCFAPLKQLMNSCNKMTPKRAMRQIQ 306
 QY 228 SLIESMSND 236
 DB 307 GELNRLACN 315

RESULT 4

S68178
 mixed-lineage protein kinase 2 (EC 2.7.1.-) - human
 C/Species: Homo sapiens (man)
 C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
 C/Accession: S68178; 138044; S32468
 R/Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps
 Eur. J. Biochem. 234, 492-500, 1995
 A>Title: Complete nucleotide sequence, expression, and chromosomal localisation of human
 A:Reference number: S68178; MUID:96128179; PMID:8536694
 A:Accession: S68178
 A:Molecule type: mRNA
 A:Residues: 1-954 <DOR>
 A:Cross-references: EMBL:X90846; NID:g971419; PIDN:CAA62351.1; PID:g971420
 R/Ratoh, M.; Hirai, M.; Sugimura, T.; Terada, M.
 Oncogene 10, 1447-1451, 1995
 A>Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase
 A:Reference number: 138044; MUID:95249256; PMID:7731697
 A:Accession: 138044
 A:Status: translated from GB/EMBL/DBJ

A:Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806, 'R', 808-817, 'A', 819-954 <RES>
 A:Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA8531.1; PID:g758593
 R/Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
 Eur. J. Biochem. 213, 701-710, 1993
 A>Title: Identification of a new family of human epithelial protein kinases containing ty
 A:Reference number: S32467; MUID:93238756; PMID:8477742
 A:Accession: S32468
 A:Molecule type: mRNA
 A:Residues: 244-464, 'AQAAGRQPHQPALML' <DO>
 C:Genetics:
 A:Gene: GDB:MLK2; GDB:MST
 A:Cross-references: GDB:362654; GDB:624810; OMIM:600137
 A:Map position: 19q13.1-19q13.2
 C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein ki
 F:23-76/Domain: SH3 homology <SH3>
 F:96-364/Domain: protein kinase homology <KIN>
 F:104-112/Region: protein kinase ATP-binding motif
 F:384-405/Region: leucine zipper motif
 F:419-440/Region: leucine zipper motif
 F:449-463/Region: basic
 F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 37.7%; Score 490.5; DB 1; Length 954;
 Best Local Similarity 42.3%; Pred. No. 5.3e-22;
 Matches 113; Conservative 35; Mismatches 92; Indels 27; Gaps 6;

QY 1 YRAKWSQDEKVAVKL-LKIEK-----EAEILSVLSHRNIQFYGVILEPPNY 48
 DB 113 YRALW--RGEVAVKARLDPEKDPVATVTAQVCOEARLFGALQHPNIIALRGACINPPLH 170
 QY 49 GYTEVASIGSLVDYINSRSEEMDMHMTWATDVAKGMYLHMEAPYKVIHRDLKSR 107
 DB 171 CLWEPFARGGALSRLVLAGRPV---PHVLVMAVQVAVAGMNYLHNDAPVPIIHRDLKSI 226
 QY 108 NVVI-----AADGVLIKIDFGASRPHNHTTMSLVGTFPMMAPEVIOQLPVSETCDT 159

```

Db      227 NIIILFAIENHNADIVLKTDFGLAREHMKTKTKMSAGTYAMAEVIRLSLFSKSSDV 286
Qy      160 YSGVVLVEMLTREVPFKLEGLOVAMLVKNERLTIPSSCPSPAEILHQCWEADAKK 219
Db      287 WSGVVLVEMLTREVPFKLEGLOVAMLVKNERLTIPSTCEPPARLIEECDDDPHG 346
Qy      220 RPSFKOIIISLESMSNDTSLPDKNSF 246
Db      347 RPDFGSILKRLKEVIEGALFQMPLESE 373

```

RESULT 5

```

JC2363
protein kinase (EC 2.7.1.37) ZPK - human
N/Alternate names: leucine-zipper protein kinase
C/Species: Homo sapiens (man)
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
C/Accession: JC2363
R/Reddy, U.R.; Pleasure, D.
Biochem. Biophys. Res. Commun. 202, 613-620, 1994
A/Title: Cloning of a novel putative protein kinase having a leucine zipper domain from
A/Reference number: JC2363; MID:94311945; PMID:8037767
A/Accession: JC2363
A/Molecule type: mRNA
A/Residues: 1-668 <RED>
A/Cross-references: EMBL:U07358
A/Experimental source: brain
A/Note: the nucleotide sequence for this amino acid sequence is inconsistent with that of
he codon ACC for residue 661 as Pro, the codon GAACACCTCTCCA for residues 664-668 as A
C/Comment: This protein belongs to the family of non-receptor kinase.
C/Genetics:
A/Gene: GDB:ZPK
A/Cross-references: GDB:383963; OMIM:600447
A/Map position: 12q13-12q13
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
F/123-371/Domain: protein kinase homology <KIN>
F/131-139/Region: protein kinase ATP-binding motif
F/143-171/Region: leucine zipper motif
F/158-145/Region: nucleotide-binding motif A (P-loop)
F/152/Active site: Lys #status predicted

```

```

Query Match      37.7%; Score 490; DB 2; Length 668;
Best Local Similarity 44.3%; Pred. No. 4, 1e-22;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

```

```

Qy      10 KEVAVKKLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNISN 67
Db      147 EEVAVKKVRDL-KETDIKRLKRLKHPNITTFKGVCTQACYCILMEFCAGOLYEVLRAG 205
Qy      68 RSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVVIADGVLIKIDFGASR- 126
Db      206 RPYTPSL--LVDSMGIAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260
Qy      127 FHNHTHMSLVGTFPPMAPEVIOSLPVSECTDYSGVVLMEMLTRVFPKLEGLOVAM 186
Db      261 LSDKSTKMSFAGTVAMAPVIRNEPVSKEVDIWSFGVLMELTGTIRPKVDSSAIIW 320
Qy      187 LVVEKNERLTIPSSCPSPAEILHQCWEADAKKPSFKOIIISLESMSNDT-SLPDK 242
Db      321 GVGSNSLHLFVPSSCPDPGFKILLRQCWNKSPRNPSFRQILLHLDIASADVLSTPGE 377

```

RESULT 6

```

A55318
serine/threonine protein kinase (EC 2.7.1.-) DLK - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-Sep-1999
C/Accession: A55318
R/Holzman, L.B.; Merritt, S.E.; Fan, G.
J. Biol. Chem. 269, 30808-30817, 1994
A/Title: Identification, molecular cloning, and characterization of dual leucine zipper

```

```

S.
A/Reference number: A55318; MID:95074107; PMID:7983011
A/Accession: A55318
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-888 <HOL>
A/Cross-references: GB:U14636; NID:9602677; PIDN:AAA57280.1; PID:9602678
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; leucine zipper; phosphotransferase
F/156-404/Domain: protein kinase homology <KIN>
F/164-172/Region: protein kinase ATP-binding motif

```

```

Query Match      37.7%; Score 488; DB 2; Length 888;
Best Local Similarity 43.9%; Pred. No. 5, 3e-22;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

```

```

Qy      10 KEVAVKKLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNISN 67
Db      180 EEVAVKKVRDL-KETDIKRLKRLKHPNITTFKGVCTQACYCILMEFCAGOLYEVLRAG 238
Qy      68 RSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVVIADGVLIKIDFGASR- 126
Db      239 RPYTPSL--LVDSMGIAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 293
Qy      127 FHNHTHMSLVGTFPPMAPEVIOSLPVSECTDYSGVVLMEMLTRVFPKLEGLOVAM 186
Db      294 LSDKSTKMSFAGTVAMAPVIRNEPVSKEVDIWSFGVLMELTGTIRPKVDSSAIIW 353
Qy      187 LVVEKNERLTIPSSCPSPAEILHQCWEADAKKPSFKOIIISLESMSNDT-SLPDK 242
Db      354 GVGSNSLHLFVPSSCPDPGFKILLRQCWNKSPRNPSFRQILLHLDIASADVLSTPGE 410

```

RESULT 7

```

JC5399
dual leucine zipper kinase (EC 2.7.-.-) - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 28-May-1999
C/Accession: JC5399
R/Matsu, N.; Sarkar, G.; Shuto, T.; Mats, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E.
Biochem. Biophys. Res. Commun. 229, 571-576, 1996
A/Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.
A/Reference number: JC5399; MID:97127443; PMID:8954939
A/Accession: JC5399
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-888 <MAT>
C/Comment: This enzyme is involved in regulating cell function in the musculoskeletal syste
C/Genetics:
A/Gene: rDLK
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; phosphotransferase
F/154-75, 89-98/Region: glycine-rich
F/156-454/Domain: kinase catalytic #status predicted <CAT>
F/156-404/Domain: protein kinase homology <KIN>
F/164-172/Region: protein kinase ATP-binding motif
F/1421-449/Region: leucine zipper motif
F/1472-500/Region: leucine zipper motif
F/557-888/Region: glycine-serine-proline rich #status predicted

```

```

Query Match      37.5%; Score 488; DB 2; Length 888;
Best Local Similarity 43.9%; Pred. No. 7e-22;
Matches 104; Conservative 44; Mismatches 79; Indels 10; Gaps 6;

```

```

Qy      10 KEVAVKKLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNISN 67
Db      180 EEVAVKKVRDL-KETDIKRLKRLKHPNITTFKGVCTQACYCILMEFCAGOLYEVLRAG 238
Qy      68 RSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVVIADGVLIKIDFGASR- 126
Db      239 RPYTPSL--LVDSMGIAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 293
Qy      127 FHNHTHMSLVGTFPPMAPEVIOSLPVSECTDYSGVVLMEMLTRVFPKLEGLOVAM 186

```

Db 294 LSKSKTKMSAGCIVAMMAPEVINEPSEKVDIMSGCVIMELLTGIEIPKVDSSAIIW 353
 QY 187 LVYKNERLTIPSSCPSPFAELHQCWEADAKKPSFKQIISLESMSNDT-SLPDK 242
 Db 354 GVGSNLHLHPVPSSCPDGFILLRQCWNSKPRRPSFRQILHLIDIASADVISTPGE 410

RESULT 8

A53800
 mixed-lineage protein kinase (EC 2.7.1.-) 3 - human
 N:Alternate names: protein kinase PTK1; protein kinase SPRK
 C:Species: Homo sapiens (man)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R/Accession: A53800; 158395
 R/Callio, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
 J. Biol. Chem. 269, 15092-15100, 1994
 A>Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
 A:Reference number: A53800; MIMD:94253068; PMID:8195146
 A:Accession: A53800
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-847 <GAL>
 A:Cross-references: GB:U07747; NID:9464027; PIDN:AAA19647.1; PID:9464028
 R/ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.
 Oncogene 9, 1745-1750, 1994
 A>Title: MLK-3: Identification of a widely expressed protein kinase bearing an SH3 domain
 A:Reference number: 158395; MIMD:94239754; PMID:8183572
 A:Accession: 158395
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-847 <RES>
 A:Cross-references: GB:U3276; NID:9488295; PIDN:AAA59859.1; PID:9488296
 C:Genes: GDB:MLK3; PTK1; SPRK
 A:Gene: GDB:MLK3; PTK1; SPRK
 A:Cross-references: GDB:134755; OMIM:600050
 A:Map position: 11913.1-11913.3
 C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
 F/48-100/Domain: SH3 homology <SH32>
 F/115-183/Domain: protein kinase homology <KIN>
 F/123-131/Region: protein kinase ATP-binding motif
 F/403-424/Region: leucine zipper motif
 F/438-459/Region: leucine zipper motif
 F/468-482/Region: basic

Query Match 37.1%; Score 482.5; DB 1; Length 847;
 Best local Similarity 41.1%; Pred. No. 1.4e-21;
 Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;

QY 1 YRAKWIISQ-----DKKAVKKLTKIEKEAELISVLSHRNIIQFYGVILEPPNYG 49
 Db 132 YRSMSGELVAVKARODPEDISV-TAESVROEARLFMLAHPNIIALKAVCLEBPNIC 190
 QY 50 IYVEVSLGSLDYINSNSEEMDMHIM-TWATDVAKGMYHMEAPYKVIHRDLKSRN 108
 Db 191 LVMEIYAGGELSRALGRVP-----PVLVNMVVOJARGMYHHCALYVPIHRDLKSRN 246
 QY 109 VV-----IADGV-----LKICDGFASRFNHTTMSLVGTFPMMAPEVIOSLPYSETCDTY 160
 Db 247 ILLQPIESDMDHKLTKIDFLAREMHTQMSAAGTAMAPVIAKSTFSKSDWA 306
 QY 161 SYGVVLMEMITRVVPFKGLEGOVAMLVYKNERLTIPSSCPSPFAELHQCWEADAKK 220
 Db 307 SFGVLLMELLTGVPVKGIDCLAVGVAVNKLTLPIPTCPRPFAOLMADCAODPHRR 366
 QY 221 PSFKQIISLES 233
 Db 367 PDFASILQOLEAL 379

RESULT 9

MAPK delta-1 protein kinase - Arabidopsis thaliana
 N:Alternate names: protein P14F18.20
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48544
 R/Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesternhoeft, A.; Bancroft, submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24490
 A:Accession: T48544
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-886 <BEV>
 A:Cross-references: EMBL:AL163812
 A:Experimental source: cultivar Columbia; BAC clone P14F18
 C:Genes:
 A:Map position: 5
 A:Intons: 148/3; 180/3; 327/3; 362/2; 568/3; 658/1; 671/3; 694/3; 728/2; 754/3; 810/3; 8
 A>Note: P14F18.20

Query Match 36.9%; Score 480; DB 2; Length 886;
 Best local Similarity 42.9%; Pred. No. 2.1e-21;
 Matches 97; Conservative 41; Mismatches 72; Indels 16; Gaps 5;

QY 1 YRAKWIISQDKKAVKKL-----KIEKEAELISVLSHRNIIQFYGVILEPPNYG 50
 Db 664 YRAEM--NCTEYAVKKFLDQDPSGALQFKSEIEMLRPNVVLFWGATREPNFSI 721
 QY 51 VTEYASLGSLDYINSNSEEMDMHIMTWATDVAKGMYHMEAPYKVIHRDLKSRNV 110
 Db 722 LTFEPLRGSLYRLH-RPHQDKEKRRMADLVAKGMNYLHTSPT-VVHRDLKSPML 779
 QY 111 IADGVLTIKCDGASRFNHT--TMSLVGTFPMMAPEVIOSLPYSETCDTYSYGVVLM 168
 Db 780 VDKMNVKVCDFGLSRMKRHTYLSKSTAGTEPMMAPEVLRPEPNEKCDVYSFGVILME 839
 QY 169 MTRREVPFKGLEGOVAMLVYKNERLTIPSSCPSPFAELHQCWE 214
 Db 840 LATSVPWKGILNPMQVGVAGVGRNRLRLPDDIDITVAQIIECQ 885

RESULT 10
 P96763
 hypothetical protein P25P22.8 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F96763
 R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; W. Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luhrs, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Iker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MIMD:21016719; PMID:11130712
 A:Accession: F96763
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1030 <STO>
 A:Cross-references: GB:AE005173; NID:96692730; PIDN:AAF24836.1; GSPDB:GN00141
 C:Genes:
 A:Gene: F25P22.8
 A:Map position: 1

Query Match 36.6%; Score 475.5; DB 2; Length 1030;
 Best local Similarity 38.3%; Pred. No. 4.4e-21;
 Matches 102; Conservative 48; Mismatches 85; Indels 31; Gaps 7;

QY 1 YRAKWIISQDKKAVKKL-----KIEKEAELISVLSHRNIIQFYGVILEPPNYG 50
 Db 763 YRGDM--HCTEYAVKKFLDQDITGEALREFRSEVIMMKLRPNIVLFWGATREPNLSI 820


```
Qy      229 ILESMSNDTSLP---DKCNSFL 247
          |      : | | | |
Db      952 ALTPLQR-LVIPAYVDQLNSRL 972
```

RESULT 14

Protein kinase homolog F6E21.90 - *Arabidopsis thaliana*
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
 C/Accession: T10671
 R/Beyan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
 submitted to the Protein Sequence Database, June 1999
 A/Reference number: Z16533
 A/Accession: T10671
 A/Molecule type: DNA
 A/Residues: 1-412 <BEV>
 A/Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
 A/Experimental source: cultivar Columbia; BAC clone F6E21
 C/Genetics:
 A/Gene: ATSP:F6E21.90
 A/Map position: 4
 A/Introns: 300/2
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 ;135-352/Domain: protein kinase homology <Kin>

RESULT 15

serine/threonine-protein kinase ctrl - Arabidopsis thaliana
N/Alternate names: protein F17C15.150
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_17-Nov-2000
C/Accession: T48400; A45178
R/Beyan, M.; Pohl, T.; Weizsaecker, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
;Reference number: Z24492

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-468,470-821 <KTE>
A>Note: sequence extracted from NCBI backbone (NCBI:P.124878)
C:Genetics:
A:Map position: 5
A:introns: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/3; 715/2
A>Note: F17C15.150
C:Superfamily: unaassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologs
C:Keywords: ATP
P:548-812/Domains: protein kinase homology <KIN>
F:557-565/Region: protein kinase ATP-binding motif

Search completed: December 5, 2003, 09:20:58
Job time : 11.2593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: December 5, 2003, 08:32:56 ; Search time 7.03764 Seconds
(without alignments)
1650.637 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277

Perfect score: 1300
Sequence: 1 YRAKMTISQKEVAVKKLTKI.....STLESMSNDTSIPDKNSFL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510.5	39.3	394	1	M3K9_HUMAN
2	490.5	37.7	954	1	M3KA_HUMAN
3	490	37.7	859	1	M3KC_HUMAN
4	488	37.5	888	1	M3KC_MOUSE
5	488	37.5	888	1	M3KC_RAT
6	442	34.0	821	1	CTRI_ARATH
7	395	30.4	579	1	M3K7_HUMAN
8	395	30.4	606	1	M3K7_HUMAN
9	369.5	28.4	1130	1	ABL1_HUMAN
10	368.5	28.3	746	1	ABL1_MLVB
11	368.5	28.3	1123	1	ABL1_MOUSE
12	365.5	28.1	1182	1	ABL1_HUMAN
13	364.5	28.0	1520	1	ABL1_MOUSE
14	355.5	27.3	1584	1	KYK1_DICDI
15	351.5	27.0	439	1	ABL1_FSVHY
16	349.5	26.9	393	1	M3K7_MOUSE
17	345	26.5	536	1	FTN_XIPHE
18	343	26.4	505	1	FRK_HUMAN
19	337.5	26.0	536	1	FTN_XENLA
20	335.5	25.8	536	1	FTN_HUMAN
21	335	25.8	819	1	FGRI_CHICK
22	335	25.8	822	1	FGRI_HUMAN
23	335	25.8	822	1	FGRI_MOUSE
24	334.5	25.7	533	1	FTN_MOUSE
25	332.5	25.6	587	1	SRC_AVIS2
26	332	25.5	517	1	FGRI_MOUSE
27	331.5	25.5	1337	1	PR2_DROME
28	330	25.4	410	1	KYK2_DICDI
29	330	25.4	822	1	FGRI_RAT
30	329.5	25.3	1224	1	ABL1_CAEL
31	329	25.3	806	1	CEK2_CHICK
32	328.5	25.3	450	1	CSK_HUMAN
33	328.5	25.3	532	1	SRC_CHICK

34	327.5	25.2	531	1	SRC1_XENLA	P13115 xenopus lae
35	327.5	25.2	531	1	SRC2_XENLA	P13116 xenopus lae
36	327	25.2	806	1	FGRI_HUMAN	P22607 homo sapien
37	326.5	25.1	557	1	SRC_AVIS1	P14085 avian sarco
38	325.5	25.0	533	1	FTN_CHICK	Q05876 gallus gall
39	324.5	25.0	533	1	SRC_HUMAN	P12931 homo sapien
40	324.5	25.0	568	1	SRC_AVIS3	P14084 avian sarco
41	324.5	25.0	801	1	FGRI_MOUSE	Q61851 mus musculu
42	324	24.9	535	1	YRK_CHICK	Q02977 gallus gall
43	323	24.8	812	1	FGRI_XENLA	P22182 xenopus lae
44	322.5	24.8	450	1	CSK_RAT	P32577 rattus norv
45	321	24.7	528	1	YES_AVISY	P00527 avian sarco

ALIGNMENTS

RESULT 1	ID	M3K9_HUMAN	STANDARD:	PRT:	394 AA.
AC	P80152;				
DT	01-FEB-1994 (Rel. 28, Created)				
DT	01-FEB-1994 (Rel. 28, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Mitogen-activated protein kinase kinase 9 (EC 2.7.1.-) (Mixed				
DE	kinase kinase 1) (Fragment).				
GN	MAP3K9 OR MLK1 OR PRKEL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Colon epithelium;				
RX	MEDLINE=93238756; PubMed=8477742;				
RA	Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;				
RT	"Identification of a new family of human epithelial protein kinases				
RT	containing two leucine/isoleucine-zipper domains."				
RL	Eur. J. Biochem. 213:701-710(1993).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF				
CC	COLONIC, BREAST AND OESOPHAGEAL ORIGIN.				
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.				
CC	MAP KINASE KINASE SUBFAMILY.				
DR	PIR; S32467; U00229.				
DR	HSSP; P12931; IFMK.				
DR	Genew; HGNC:6861; MAP3K9.				
DR	MIM; 600136; -				
DR	GO; GO:0005524; F:ATP binding activity; NAS.				
DR	GO; GO:0004708; F:MAP kinase activity; NAS.				
DR	GO; GO:0006468; F:protein amino acid phosphorylation; NAS.				
DR	InterPro; IPR000719; Prot_Kinase.				
DR	InterPro; IPR002290; Ser_Thr_kinase.				
DR	Pfam; PF00069; pkinase; 1.				
DR	PRINTS; PR00109; TYRKINASE.				
DR	ProDom; PD000001; Prot_kinase; 1.				
DR	SMART; SM00219; Tyrc; 1.				
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.				
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.				
DR	PROSITE; PS00111; PROTEIN KINASE DOM; 1.				
KW	Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;				
KW	ATP-binding.				
FT	NON_TER	1			
FT	DOMAIN	3	271		
FT	NP_BIND	9	17		
FT	BRINDING	30	30		
FT	ACT_SITE	127	127		
FT	DOMAIN	289	310		
FT	DOMAIN	324	345		
FT	DOMAIN	354	368		
SQ	SEQUENCE	394 AA;	44975 MW;		
Query Match		39.3%;	Score 510.5;		
			DB 1;		
			Length 394;		

Best Local Similarity 42.7%; Pred. No. 1,6e-34;
Matches 109; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAKWSIDKEVAVK-----KLIKIEKEAELISVLSHRNIIQFVGLLEPPNY 48
DB 18 YRAFWIGD--EVAVKAARHDPDDISQITBNVQEAFLFAMLGHPNIALRGVCLKEPNL 75
QY 49 GIVTEYASLSGLVDYINSNSEMDMDHITATDVAKGKHVYLMKAPVYVTHRDLSKR 108
DB 76 CLWMEFARGSPILRVLSGKR--IPDILVMAVQVARGNNYLHDEALVPIIHRDLKSSN 132
QY 109 VVI-----AADGVLCICDFGASRPHNHTMSLVGTFPMAPEVYIOSLPVSETCDY 160
DB 133 ILLQKVENDDLSNKILIKIDFGIARPMHRTTQSAAGTAMAPVIRASMSKSDVW 192
QY 161 SYGVVLEMTLREVPFPGKLGLOVAVLVKERNLTIPTSCPSFALLHQCWEADAKR 220
DB 193 SYGVLLMELLTGVEPFGIDGLRVAGVANMKLALPIPTCPBPFAKLMEDCWPDPHSR 252
QY 221 PSEFKQIISLES 233
DB 253 PSEFTNILDQLTTI 265

RESULT 2
M3KA_HUMAN STANDARD; PRT; 954 AA.
ID Q02779; Q12761; Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
GN MAP3K10 OR MLK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RT "Complete nucleotide sequence, expression, and chromosomal
localisation of human mixed-lineage kinase 2.";
RT Eur. J. Biochem. 234:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=85249256; PubMed=7731697;
RA Katoh M., Hirai M., Sugimura T., Terada M.,
RT "Cloning and characterization of MST, a novel (putative)
serine/threonine kinase with SH3 domain.";
RT Oncogene 10:1447-1451(1995).
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.,
RT "Identification of a new family of human epithelial protein kinases
containing two leucine/isoleucine-zipper domains.";
RT Eur. J. Biochem. 213:701-710(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X90846; CAA62351.1; -
DR EMBL; Z48615; CAA8531.1; -
DR PIR; S68178; S68178.
DR HSSP; P13362; 1FGK.
DR Genew; HGNC:6849; MAP3K10.
DR MIM; 600137; -
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR GO; GO:007254; P:JNK cascade; TAS.
DR GO; GO:007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002250; Ser_thr_kinase.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50002; SH3; 1.
KM Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; SH3 domain.
FT DOMAIN 2
FT 16 81
FT DOMAIN 98 360
FT NP_BIND 104 112
FT BINDING 125 125
FT ACT_SITE 222 222
FT DOMAIN 384 405
FT DOMAIN 419 440
FT DOMAIN 449 463
FT DOMAIN 462 464
FT CONFLICT 465 480
FT CONFLICT 471 471
FT CONFLICT 807 807
FT CONFLICT 818 818
SQ SEQUENCE 954 AA; 103623 MW; 538F4AA559B0ABA CRC64;
Query Match 37.7%; Score 490.5; DB 1; Length 954;
Best Local Similarity 42.3%; Pred. No. 2e-32;
Matches 113; Conservative 35; Mismatches 92; Indels 27; Gaps 6;

QY 1 YRAKWSIDKEVAVK--LKIEK-----EAEILSVLSHRNIIQFVGLLEPPNY 48
DB 113 YRALM--RGEVAVKAARLDPEKDPVTAEGVCGEARLFGALQHNNIIALGACLNPPHL 170
QY 49 GIVTEYASLSGLVDYINSNSEMDMDHIM-TWATDVAKGKHVYLMKAPVYVTHRDLSKR 107
DB 171 CLWMEFARGSPILRVLSGKR--IPDILVMAVQVARGNNYLHDEALVPIIHRDLKSSI 226
QY 108 NVVI-----AADGVLCICDFGASRPHNHTMSLVGTFPMAPEVYIOSLPVSETCDY 159
DB 227 NLLIEAIEAHNHLATVLTIKIDFGIARPMHRTTQSAAGTAMAPVIRASMSKSDV 286
QY 160 SYGVVLEMTLREVPFPGKLGLOVAVLVKERNLTIPTSCPSFALLHQCWEADAKR 219
DB 287 WSFGVLLMELLTGVEPFGIDGLRVAGVANMKLALPIPTCPBPFAKLMEDCWPDPHSR 346
QY 220 PSEFKQIISLESMSNDTSLPDKNSF 246
DB 347 RDPFGSILKRLEVEIQSALFQWPLESF 373

RESULT 3

M3KC_HUMAN
ID M3KC_HUMAN STANDARD; PRT; 859 AA.
AC 012852;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK).
GN MAP3K12 OR ZPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Teratocarcinoma;
RA MEDLINE=94311945; PubMed=8037767;
RX Reddy U.R., Pleasure D.;
RT "Cloning of a novel putative protein kinase having a leucine zipper
RT domain from human brain."
RT Biochem Biophys Res Commun. 202:613-620(1994).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07358; AAA67343.1; -.
DR HSSP: P12931; IFMK.
DR Genew: HGNC:6851; MAP3K12.
DR MIM: 600447; -.
DR GO: GO:0005737; Cytoplasm; TAS.
DR GO: GO:0005886; Cytoplasmic membrane; TAS.
DR GO: GO:0007254; P:JNK cascade; TAS.
DR InterPro: IPR000719; Prot_Kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_Kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP, FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 125 366 PROTEIN KINASE.
FT NP_BIND 131 133 ATP (By SIMILARITY).
FT BINDING 152 152 ATP (By SIMILARITY).
FT ACT_SITE 236 236 BY SIMILARITY.
FT DOMAIN 666 666 POLY-PRO.
FT DOMAIN 720 725 POLY-GLU.
SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;
Query Match 37.7%; Score 490; DB 1; Length 859;
Best Local Similarity 44.3%; Pred. No. 1.9e-32;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEAVAKLILKIEKAEI--LSVLSHRNITQFYGVILLEPPNYGIVTEVASTGLSYDYNIN 67
Db 147 EEVAAYKARDL-KETDILKRLKLPNITTFGVCTQACVCLLWECAGQGVYEVLRAG 205
QY 68 REEMDMDHIMTWADVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLTAKICDFGASR- 126
Db 206 RVTYPSL-LVDMSGINGAGMYVLIH---KIHHDLKSPNNLITVDVVKISDPSTKAE 260
QY 127 FNNHTHSLVGTFFPMADVEVIQSLPVSETCTYSYGVVLMELTREVPFKGLEQIVAM 186
Db 261 LBDKSTKMSFAGTVAMMAPEVIRNPEVSEKVDIMSGVLMELNGEIPYKVDSSAIIW 320
QY 187 LVENKNERLTSSCPSPFAELHOCWEADAKRSPFOIILSEMSYDT-SLPXK 242
Db 321 GVGNSLHLVPSSCPDGFKILLRCQWNSKPNRRSPFQILHLHDIASADVLTSPGE 377
RESULT 4
M3KC_MOUSE
ID M3KC_MOUSE STANDARD; PRT; 888 AA.
AC 060700; P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing
DE kinase) (DLK).
GN MAP3K12 OR ZPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Brain;
RX MEDLINE=95074107; PubMed=7983011;
RA Holzman L.B., Merritt S.E., Fan G.;
RT "Identification, molecular cloning, and characterization of dual
RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
RT that defines a second subfamily of mixed lineage kinases."
RL J. Biol. Chem. 269:30808-30817(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
RX MEDLINE=96365388; PubMed=8769565;
RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
RT "Cell-specific expression of the ZPK gene in adult mouse tissues."
RL DNA Cell Biol. 15:631-642(1996).
RN [3]
RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX MEDLINE=96279269; PubMed=8663324;
RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
RT "Characterization of dual leucine zipper-bearing kinase, a mixed
RT lineage kinase present in synaptic terminals whose phosphorylation
RT state is regulated by membrane depolarization via calcineurin."
RL J. Biol. Chem. 271:16888-16896(1996).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
CC testis, gastrointestinal tract, stomach, liver and pancreas.
CC Within the nervous system, predominantly expressed in neurons and
CC enriched in synaptic terminals.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

	CC	Eumkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Scturognathii; Muridae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	Medline=9622609; Pubmed=8637721;	
RA	Hiral S., Izawa M., Osada S., Spyrou G., Ono S.?	
RT	"Activation of the JNK pathway by distantly related protein kinases,	
RL	MEKK and MUK.";	
RM	Oncogene 12;641-650(1996).	
CC	-1- FUNCTION: May be an activator of the JNK/SAPK pathway.	
CC	Phosphorylates beta-casein, histone 1 and myelin basic protein in	
CC	vivo.	
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.	
CC	-1- COFACTOR: Magnesium.	
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By	
CC	similarity).	
CC	-1- PPM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol	
CC	under basal conditions and dephosphorylated when membrane-	
CC	associated (by similarity).	
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.	
CC	MAP KINASE KINASE KINASE SUPERFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcement/	
CC	or send an email to license@ebi.ac.uk).	
CC	-----	
DR	EMBL; D49785; BAA08621.1; --	
DR	HSSP; P12931; IFMK.	
DR	InterPro; IPRO00719; Prot_kinase.	
DR	InterPro; IPRO02290; Ser_Thr_pkinase.	
DR	InterPro; IPRO01245; Tyr_pkinase.	
DR	Pfam; PF00069; pkinase; 1	
DR	PRINTS; PR00109; TYRKINASE.	
DR	ProDom; PD000001; Prot_kinase; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;	
KW	Phosphorylation; Magnesium; Membrane.	
FT	DOMAIN 158 399 PROTEIN_KINASE.	
FT	NP_BIND 164 172 ATP (BY SIMILARITY).	
FT	BINDING 185 185 ATP (BY SIMILARITY).	
FT	ACT_SITE 269 269 BY SIMILARITY.	
FT	DOMAIN 56 62 POLY-GLY.	
FT	DOMAIN 668 671 POLY-PRO.	
FT	DOMAIN 698 701 POLY-PRO.	
FT	DOMAIN 753 758 POLY-GLU.	
SQ	SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;	

	Query Match 37.5%; Score 488; DB 1; Length 888;	
	Best Local Similarity 44.3%; Pred. No. 2; se-33;	
	Matches 105; Conservative 42; Mismatches 80; Indels 10; Gaps 6;	
QY	10 KEVAVVKLLTKIEKEAEI--LSVLSHRIIOFYGVILEPPNYGIVTEYASLSGYDIYN 67	
Dd	180 EEVAVKVRDPL-KETDKIKLRKLKHPILITFGKCTGPACYCIMFCAGQLYEVRAG 238	
QY	68 RSEMDMDHMTATVDAKGMHYLMHEAPKVHIRLDKSNVVYADGVULKCDPGASR- 126	
Dd	239 RPTVPSL--LVDMSMGIAGSMNVLIHLN--KIHRDLKSNNMLITDYDVVKISDFGSKE 293	
QY	127 FHNHTTHMSLVGFPPMNAPEVIQSLPVSEFTCDTYSYGCVILMENLTREVPFKGLEGOVAN 186	
Dd	294 LSKSTKMSFAEGVAVAMAPFEVINPESEKVIDWSFGVILMELLTGEIPYKDVSASAIIW 353	
QY	187 LVVEKNERLTIPISSCRSPFAELHQCCEAIAKKRPRPKOIIISLEMSNT-SLPDK 242	
Dd	354 GVGSNSLAHLVPSSCPDGFKILLRQCENRRKRPRRSRQILLHLDIASADVLPDP 410	

RESULT 6
 CTR1 ARATH STANDARD; PRT; 821 AA.
 ID CTR1 ARATH
 AC 005609;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
 GN CTR1 OR AT5G03730 OR F17C15_150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosid II; Brassicales; Brassicaceae; Arabidopsids.
 OK NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Seedling;
 RX MEDLINE=93161417; PubMed=8431946;
 RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Becker J.R.,
 RT CTR1, a negative regulator of the ethylene response pathway in
 RL Arabidopsis, encodes a member of the raf family of protein kinases,"
 RL Cell 72:427-441 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneo T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naro K., Okumura S., Shimizu S., Takeuchi C., Mada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide C.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston J.,
 RA Haberman K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoecklin T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter B., Cordum H., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum B., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Marienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Eutlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Rampeger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirsee W., Moosman P., Klein Lankhorst R.,
 RA Weitzneger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gielens J., Ardiles W.,
 RA Bens O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schott H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana.";
 RT Nature 408:823-826 (2000).
 RL Nature 408:823-826 (2000).
 CC - FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
 CC PATHWAY.
 CC - CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC - TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
 CC - MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES.
 CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
 CC COTYLEDON GROWTH IS IMPAIRED.
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MIM/RAP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L08789; AAA32779.1; -
 DR EMBL: L08790; AAA32780.1; -
 DR EMBL: A1162506; CAB82938.1; -
 DR PIR: T48400; T48400.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR002290; Ser. Thr. Kinase.
 DR InterPro: IPR001245; Tyr. Kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR ProDom: PD000001; Prot. Kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferrase; Serine/threonine-protein Kinase; ATP-binding.
 FT DOMAIN 65 69
 FT POLY-GLY.
 FT DOMAIN 135 141
 FT POLY-GLY.
 FT DOMAIN 551 809
 FT PROTEIN KINASE.
 FT NP_BIND 557 565
 FT ATP (BY SIMILARITY).
 FT BINDING 578 578
 FT ATP (BY SIMILARITY).
 FT ACT_SITE 676 676
 FT BY SIMILARITY.
 FT MUTAGEN 596 596
 FT E->K: IN CTR1-4; EXHIBITS ETHYLENE-
 FT TREATED PHENOTYPE.
 FT D->E: IN CTR1-1; EXHIBITS ETHYLENE-
 FT TREATED PHENOTYPE.
 SQ SEQUENCE 821 AA; 90306 MW; 2922DD3DCDC15BC CRC64;
 Query Match 34.0%; Score 442; DB 1; Length 821;
 Best Local Similarity 40.2%; Pred. No. 1.5e-28;
 Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;
 QY 1 YAAKISQDKVAVVKKLIE-----KEALISLVSRLNIQFGVLEPPNYGI 50
 DB 566 HRAEHGSD--AAVITLMEQDHAARVNEFLAEVAIMKRLRHPNIVLFNGAVTQPPNLSI 623
 QY 51 VEEVYSLGSLDYIN-SNRSEMDMDHMTATDYAKGMHYHMEAPVYVIRDLKSRV 109
 DB 624 VTEVYSRGSLLYLKSGAREQDERRRLSMAYDAKGMVYHNRP-PIVRDLKSPYL 682
 QY 110 VIADQVLKICDGFSGARFHNH--TMSLVGTFPMMAPEVIOSLPVSETCDTYSYGVVLW 167
 DB 683 LVDKKYTVVCCFGLSRKASTFPLSSKAAGTPENMAPEVLDEPSNEKSDVYSFVILW 742
 QY 168 EMLTEVPVKKGLGQVAMLVKNERLTTPSCRSFELLHQCEADAKRPPKQII 227
 DB 743 ELATIQPWNINPAQVAAVGFCKRLIEPNNINPOVAAITGCTWEPWRRPSPATIM 802
 QY 228 SLTESN 233
 DB 803 DLRLP 808
 RESULT 7
 M3K7 MOUSE
 ID M3K7 MOUSE STANDARD; PRT; 579 AA.
 AC Q62073;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase 7 (BC 2.7.1.-)
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
 DE activated kinase 1).
 GN MAP3K7 OR TAK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96123277; PubMed=8533096;
 RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
 RA Taniguchi T., Nishida E., Matsumoto K.,
 RT "Identification of a member of the MAPKK family as a potential

```

RT mediator of TGF-beta signal transduction.";
RL Science 270:2008-2011(1995).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D76446; BAA1184.1; -.
DR HSSP: P08631; IAD5.
DR MGD; MGI:1346877; Map3k7.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 8 16 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT BINDING 42 50 ATP (BY SIMILARITY).
FT ACT SITE 63 63 ATP (BY SIMILARITY).
FT ACT SITE 156 156 BY SIMILARITY.
SQ SEQUENCE 579 AA; 64227 MW; 97C8F6FC8E283EE CRC64;

Query Match 30.4%; Score 395; DB 1; Length 579;
Best Local Similarity 37.1%; Pred. No. 6; 9e-25;
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISODKEVAVKLLKIKKEAE-----ILSVLSHRNIIQFYGVILPEPNYGIWT 52
DB 52 KAKW--RAKVAK--QISESEKRAFYELQSLRVNHPNIVKLYGACLN--VCLWA 104
QY 53 EVASLGSLYDYINSNSEN--DMDHIMTWADVAKGMHYLHMEAPVKYIHRDLKSRV 109
DB 105 EYAEAGSLVYVHLHG--AEPLPYTAHAWMSWCLQSGVAYLHSMQPKALIHRLKPNL 162
QY 110 VIADG-VLKICPFG-ASRHHNTTHMSLVGTFRMAPEVYQSLPVSETCDTYSYGVLM 167
DB 163 LTVAGGTVLKICFGTACDIQTMNTNK--GSAWMAPEVEFSNSSEKDVSWGIILM 220
QY 168 EMVTRVPEPKGLSG--LQVAVLVEKKERLTIISSCPSPAEILLHOCWEADAKRPSFK 225
DB 221 EVITRRKPFPEIGSPAFRIW-AVHNGTRPPLIKNPKPIESLMTKWSKDPQRSMEE 279
QY 226 IISIL 230
DB 280 IVKIM 284

RESULT 8
M3K7 HUMAN STANDARD; PRT; 606 AA.
AC Q43318; Q43317; Q43319;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).
RC TISSUE=Lung;
RX MEDLINE=98153801; PubMed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1B;
CC IsoId=Q43318-1; Sequence=Displayed;
CC Name=1A;
CC IsoId=Q43318-2; Sequence=VSP_004886;
CC Name=1C;
CC IsoId=Q43318-3; Sequence=VSP_004887, VSP_004888;
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP
CC KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB009357; BAA25026.1; -.
DR EMBL: AB009356; BAA25025.1; -.
DR EMBL: AB009358; BAA25027.2; -.
DR FIR; JC5955; JC5955.
DR FIR; JC5956; JC5956.
DR HSSP; P08631; IAD5.
DR Genew; HGNC:6859; MAP3K7.
DR MIM: 602614; -.
DR GO; GO:0004709; F:MAP kinase kinase activity; TAS.
DR GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing;
FT DOMAIN 8 14 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT BINDING 42 50 ATP (BY SIMILARITY).
FT ACT SITE 63 63 ATP (BY SIMILARITY).
FT ACT SITE 156 156 BY SIMILARITY.
FT VARSPLIC 404 430 Missing (in isoform 1A).
FT VARSPLIC 509 518 /FTID=VSP_004886.
FT VARSPLIC 519 606 /FTID=VSP_004887.
FT VARSPLIC 519 606 /FTID=VSP_004888.
FT VARSPLIC 519 606 /FTID=VSP_004889.
SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 30.4%; Score 395; DB 1; Length 606;
Best Local Similarity 37.1%; Pred. No. 7; 3e-25;
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

```

QY 2 RAKWISQDKEAVKLLKIEKEA-----ILSVLSHRNIIQFYGVILPPNYGIWT 52
 DB 52 KAKM--RADVAIK--QIESESEKAFIVELRQLSRVNHPNIVKIXGACINP--VCLWM 104
 QY 53 EVASIGSLVDYINSNSEEM--DMDHMTWATDVAKGMYLMEAPVVKYHDLKSRV 109
 DB 105 EVAEGGSLYNVLHG--AEPLPYTAHAAMSWCLQSGQVAYLHSMQPKALIHRLDKPPL 162
 QY 110 VIAADG-VKICDFG-ASRFNHTTMSLVGTFPMMAPEVIGSLPVSEICDTYSGVIM 167
 DB 163 LTVAGGTIVKICDFGTACDIQTHMTNNK--GSAAMMAPEVEGSSNNSEKCDVFSWGIILM 220
 QY 168 EMUTREVPKGLGEG--LOYAVMLVEKNERTLTPSSCPSPFALLHQCEWADAKRRSPFO 225
 DB 221 EVITRKPPDEIGGPAFRIM-AVHNGTRPPLIKNPKPIESLMTCTMSKDSQPSMEE 279
 QY 226 IISIL 230
 DB 280 IVKIM 284

RESULT 9
 ABL1 HUMAN STANDARD; PRT; 1130 AA.
 ID ABL1 HUMAN P00519; Q13869; Q13870; Q16133;
 AC P00519; Q13869; Q13870; Q16133;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150) (c-ABL).
 GN ABL1 OR ABL OR JTK7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 (1)
 RN RN SEQUENCE FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=90082420; PubMed=2687768;
 RA Fainstein E., Einat M., Gokkel E., Marcelle C., Croce C.M.,
 RA Gale R.P., Canaan E.;
 RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs.";
 RL Oncogene 4:1477-1481(1989).
 (2)
 RN RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RP MEDLINE=87028219; PubMed=3021337;
 RA Shitvelan E., Lifshitz B., Gale R.P., Roe B.A., Canaan E.;
 RT "Alternative splicing of RNAs transcribed from the human abl gene and from the bcr-abl fused gene.";
 RL Cell 47:277-284(1986).
 (3)
 RN RN SEQUENCE FROM N.A. (ISOFORMS IA AND IB).
 RC TISSUE=Lung carcinoma;
 RX MEDLINE=95394474; PubMed=766185;
 RA Chisoe S.W., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jia L., Ma Y.,
 RA McLary H.-J., Pan H.-Q., Sathan O.H., Toch S., Wang Z., Zhang G.,
 RA Heisterkamp N., Groffen J., Roe B.A.;
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and RT regions involved in the Philadelphia chromosome translocation.";
 RL Genomics 27:67-82(1995).
 (4)
 RN RN SEQUENCE OF 360-426 FROM N.A.
 RP MEDLINE=83245023; PubMed=6191223;
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
 RT "Homology between phosphotyrosine acceptor site of human c-abl and RT viral oncogene products";
 RL Nature 304:167-169(1983).
 (5)
 RN RN SEQUENCE OF 27-40 FROM N.A.
 RP MEDLINE=88065859; PubMed=2825022;
 RA Fainstein E., Marcelle C., Rosner A., Canaan E., Gale R.P.,
 RA Drazan O., Smith S.D., Croce C.M.;

RT "A new fused transcript in Philadelphia chromosome positive acute lymphocytic leukemia";
 RT Nature 330:386-388(1987).
 (6)
 RN RN SEQUENCE OF 825-845 FROM N.A.
 RP MEDLINE=94142331; PubMed=7545908;
 RA Inokuchi K., Futaki M., Dan K., Nomura T.;
 RT "Sequence analysis of the mutation at codon 834 and the sequence variation of codon 837 of c-abl gene.";
 RL Leukemia 8:343-344(1994).
 (7)
 RN RN STRUCTURE BY NMR OF SH2 DOMAIN.
 RP MEDLINE=92370689; PubMed=1505033;
 RA Overduin M., Rios C.B., Mayer B.J., Baltimore D.;
 RT "Three-dimensional solution structure of the src homology 2 domain of c-abl.";
 RL Cell 70:697-704(1992).
 (8)
 RN RN STRUCTURE BY NMR OF SH2 DOMAIN.
 RP MEDLINE=93101588; PubMed=1281542;
 RX Overduin M., Mayer B.J., Rios C.B., Baltimore D.;
 RA "Secondary structure of Src homology 2 domain of c-Abl by heteronuclear NMR spectroscopy in solution.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677(1992).
 (9)
 RN RN STRUCTURE BY NMR OF SH3 DOMAIN.
 RP MEDLINE=96131878; PubMed=8590002;
 RX Gosses Y.Q., Zheng J., Overduin M., Mayer B.J., Cowburn D.;
 RA "The solution structure of Abl SH3, and its relationship to SH2 in the SH(32) construct.";
 RT Structure 3:1075-1086(1995).
 (10)
 RN RN 3D-STRUCTURE MODELING OF SH3 DOMAIN.
 RP MEDLINE=95199229; PubMed=7892170;
 RX Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;
 RA "Homology modeling of the Abl-SH3 domain.";
 RL Proteins 20:203-215(1994).
 (11)
 RN RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.
 RP MEDLINE=96398698; PubMed=8805596;
 RX Nam H.-J., Haer W.G., Roberts T.M., Frederick C.A.;
 RA "Intramolecular interactions of the regulatory domains of the Bcr-Abl kinase reveal a novel control mechanism.";
 RL Structure 4:1105-1114(1996).
 (12)
 RN RN X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.
 RP MEDLINE=98365516; PubMed=9698566;
 RX Pisabarro M.T., Serrano L., Wilmanns M.;
 RA "Crystal structure of the abl-SH3 domain complexed with a designed high-affinity peptide ligand: implications for SH3-ligand interactions.";
 RT J. Mol. Biol. 281:513-521(1998).
 (13)
 RN RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 223-515.
 RP MEDLINE=20446271; PubMed=10988075;
 RX Schindler T., Bornmann W., Pellicena P., Miller W.T., Clarkson B., Kuriyan J.;
 RA "Structural mechanism for Src-571 inhibition of abelson tyrosine kinase";
 RT Science 289:1938-1942(2000).
 RL Science 289:1938-1942(2000).
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=IA;
 CC IsoId=P00519-1; Sequence=Displayed;
 CC Name=IB;
 CC IsoId=P00519-2; Sequence=VSP_004957;
 CC -I- TISSUE SPECIFICITY: Widely expressed.
 CC -I- DISEASE: PARTICIPATES IN A T(9;22)(Q34;Q11) CHROMOSOMAL TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND

CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).
 CC - SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
 CC SUBFAMILY.
 CC - SIMILARITY: Contains 1 SH2 domain.
 CC - SIMILARITY: Contains 1 SH3 domain.
 CC - DATABASE: NAME=atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chromocancer/Genes/ABL.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M16416; CA434438.1; -;
 DR EMBL; M14752; AAA51561.1; -;
 DR EMBL; U07563; AAB60394.1; -;
 DR EMBL; U07563; AAB60393.1; -;
 DR EMBL; U07561; AAB60393.1; JOINED.
 DR EMBL; S69223; AAD14034.1; -;
 DR PIR; S08519; TVHUA.
 DR PDB; 1AB2; 31-JAN-94.
 DR PDB; 2ABL; 04-SEP-97.
 DR PDB; 1AM0; 28-JAN-98.
 DR PDB; 1BB2; 25-NOV-98.
 DR PDB; 1FPU; 20-SEP-00.
 DR PDB; 1ABJ; 01-NOV-94.
 DR PDB; 1JUS; 06-NOV-02.
 DR Genew; HGNC:76; ABL1.
 DR MIM; 189980; -;
 DR GO; GO:0004713; F:Protein tyrosine kinase activity; TAS.
 DR GO; GO:0008630; P:Induction of apoptosis by DNA damage; TAS.
 DR GO; GO:0006298; P:mismatch repair; TAS.
 DR GO; GO:0007048; P:oncogenesis; TAS.
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR GO; GO:0000115; P:s-specific transcription in mitotic cell cycle; TAS.
 DR InterPro; IPR000719; Proc_kinase.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Proc_kinase; 1.
 DR ProDom; PD000093; SH2; 1.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 1.
 DR Transfaser; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
 KM Phosphorylation; SH2 domain; SH3 domain; Chromosomal translocation;
 KM 3D-structure; Alternative splicing.
 FT DOMAIN 61 121
 FT DOMAIN 127 217
 FT DOMAIN 242 493
 FT DOMAIN 605 609
 FT DOMAIN 18 22
 FT DOMAIN 605 609
 FT DOMAIN 782 1019
 FT DOMAIN 897 903
 FT SITE 26 27
 ABL ONCOGENE.

FT NP_BIND 248 256 ATP (BY SIMILARITY).
 FT FT_BINDING 271 271 ATP (BY SIMILARITY).
 FT ACT_SITE 363 363 BY SIMILARITY.
 FT MOD_RES 393 393 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 Query Match 28.4%; Score 369.5; DB 1; Length 1130;
 Best Local Similarity 32.4%; Pred. No. 1.9e-22;
 Matches 82; Conservative 55; Mismatches 101; Indels 15; Gaps 7;
 QY 1 YAAKISODKEVAVKL---LKIF---KEAIIISLSHRNIIQYGVILPEPNYGIYVE 53
 DB 257 YGVWKKYSILVAIVTKLDEDTWEVEEPLKEAAVMEIKIPNVQLIGVCTREPPYIITE 316
 QY 54 VASLSLDYINISNSSEEMDHIMTATFDVAKGKHVYLMFAVPVYIHRDLKSRVVIAA 113
 DB 317 FMTYGNLDYLEKCRQEVNAVLLVMYMTQISSANEYERK--NFIRHDLARNCLVGE 373
 QY 114 DGVLRICDPGASRFPHHTTHMSLVGT-PP--WMADEVISLPVSECTDYSGVYLWEML 170
 DB 374 NMLVAVADPGLSRMTGDTYTHAGAKPPIKKTAPESLAYNKFSIKSDVMARGLLMEIA 433
 QY 171 TREV-PFKGLEGLQVAMLVVERKNERLTIPSSCPSPFAELHOCWADAKRPSFOIISI 229
 DB 434 TYGMSPPYPGIDISQY-YELLERDYMERBECPEKRYEILMRACWQWNSDRPSFAIHQA 492
 QY 230 LESMSNDTSLPDK 242
 DB 493 FETMFOESSISDE 505
 RESULT 10
 ABL_MLVAB STANDARD; PRT; 746 AA.
 ID ABL_MLVAB
 AC P00521;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).
 GN V-ABL.
 OS Abelson murine leukemia virus.
 OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_Taxid=11788;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83221648; PubMed=6304726;
 RA Reddy E.P., Smith M.J., Srinivasan A.;
 RT "Nucleotide sequence of Abelson murine leukemia virus genome:
 RT structural similarity of its transforming gene product to other onc
 RT gene products with tyrosine-specific kinase activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
 RN [2]
 RP REVISIONS TO 588-746.
 RA Reddy E.P., Smith M.J., Srinivasan A.;
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7372-7372(1983).
 RN [3]
 RP SEQUENCE OF 233-377 FROM N.A.
 RX MEDLINE=83245023; PubMed=6191223;
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
 RT "Homology between phosphotyrosine acceptor site of human c-abl and
 RT viral oncogene products.";
 RL Nature 304:167-169(1983).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
 CC POLYPROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL; V01541; -; NOT ANNOTATED_CDS.
CC
CC DR EMBL; K00010; AAA46470.1; -.
CC DR HSSP; P00519; 2ABL.
CC DR InterPro; IPR0000719; Prot_Kinase.
CC DR InterPro; IPR000980; SH2_Kinase.
CC DR InterPro; IPR001245; Tyr_kinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR Pfam; PF00017; SH2; 1.
CC DR PRINTS; PRO0401; SH2DOMAIN.
CC DR PRINTS; PRO0109; TYRKINASE.
CC DR ProDom; PD000001; Prot_Kinase; 1.
CC DR ProDom; PD000093; SH2; 1.
CC DR SMART; SM00252; SH2; 1.
CC DR SMART; SM00219; Tyrc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS50001; SH2; 1.
CC KW Polypeptide; Tyrosine-protein kinase; Transferase; Oncogene;
CC SH2 domain.
CC FT DOMAIN 13 103 SH2.
CC FT DOMAIN 128 379 PROTEIN_KINASE.
CC SQ SEQUENCE 746 AA; 81872 MW; B9072PFF5FE9257 CRC64;
CC
CC Query Match 28.3%; Score 368.5; DB 1; Length 746;
CC Best Local Similarity 32.4%; PId No. 1.4e-22;
CC Matches 82; Conservative 54; Mismatches 102; Indels 15; Gaps 7
CC
CC QY 1 YRAKMSIDKKEAVAKLT---LKIE--KEAEILSVLSHRNIIQFYGVILPEPNYGIATVE 53
CC DB 143 YEGWKKKSLVVAVKLTKEIDMEVEEFLKEAAMKEIGHPLVQLLGGVCTREPPYIITE 202
CC
CC QY 54 YASIGSLDYDINSNSREEMDHDITWTATDVAKGMHYHMEAPVKYVTHRDLKSRNVITAA 113
CC DB 203 FMTYGNLLDYDIRECNCREVSAAVLLVMATQISSAMEYLEKK--NFIRHDLAARNCVLGE 259
CC
CC QY 114 DGVLIKCPGASRFPHNHTTHSLVGT-PP--WMAPEVIGSLPVSEPTCDYSGVGLMEML 170
CC DB 260 NHALKVAADFGLSRLMTGDTYTAHAGAKFPFKWTAPESLAYNKFESIKSDVMAGVLMETIA 319
CC
CC QY 171 TREV-PFKGLDELQVAMLVMEKNERLTISSCPRSFAELLHQWEADAKKSPFFKOITSI 229
CC DB 320 TYGMSFYGIDLSQY-YELLKDYRMERNPECGCEKVEYVELMARCAQMGNSDRPSFAIHQA 378
CC
CC QY 230 LESMSNDTSLPDK 242
CC DB 379 FELTFQESSISDE 391
CC
CC RESULT 11
CC ABL1_MOUSE STANDARD; PRT; 1123 AA.
CC AC P00520; P97896; Q61252; Q61253; Q61254; Q61255; Q61256; Q61257;
CC AC Q61258; Q61259; Q61260; Q61261;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)
CC (c-ABL).
CC GN ABL1 OR ABL.
CC OS Mus musculus (Mouse).
CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_Taxid=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC Tissue=Testis;
CC FA MEDLINE=86068561; PubMed=3317402;
CC Opi C.; Shore S.K.; Reddy E.P.;

```

"Nucleotide sequence of testis-derived c-abl cDNAs: implications for
 testis-specific transcription and abl oncogene activation.";
 Proc. Natl. Acad. Sci. U.S.A. 84:8200-8204(1987).
 [2].
 RN SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I; II; III AND IV).
 RP MEDLINE=95394474; Pubmed=7665185;
 RX Chinese S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,
 RA McLaury H.-U., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,
 RA Heisterkamp N., Groffen J., Roe B.A.;
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and
 RT regions involved in the Philadelphia chromosomal translocation.";
 RL Genomics 27:67-82(1995).
 RN [3].
 RP SEQUENCE OF 85-182 FROM N.A.
 RX MEDLINE=84106840; Pubmed=6319018;
 RA Wang J.Y.J., Ledley F., Goff S., Lee R., Groner Y., Baltimore D.;
 RT "The mouse c-abl locus: molecular cloning and characterization.";
 RL Cell 36:349-356(1984).
 RN [4].
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=88202920; Pubmed=3283651;
 RA Bernards A., Paskind M., Baltimore D.;
 RT "Four murine c-abl mRNAs arise by usage of two transcriptional
 RT promoters and alternative splicing.";
 RL Oncogene 2:297-304(1988).
 RN [5].
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 61-121.
 RX MEDLINE=95393198; Pubmed=7664083;
 RA Musacchio A., Saraste M., Wilmanns M.;
 RT "High-resolution crystal structures of tyrosine kinase SH3 domains
 RT complexed with proline-rich peptides.";
 RL Nat. Struct. Biol. 1:546-551(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASM;C; THE MYRISTOYLATED C-ABL PROTEIN
 CC WAS REPORTED TO BE NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=I;
 CC IsoId=P00520-1; Sequence=Displayed;
 CC Name=II;
 CC IsoId=P00520-2; Sequence=VSP_004959;
 CC Name=III;
 CC IsoId=P00520-3; Sequence=VSP_004958;
 CC Name=IV;
 CC IsoId=P00520-4; Sequence=VSP_004960;
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02995; AAA88241.1; -;
 DR EMBL; U14721; AAB60451.1; -;
 DR EMBL; U14720; AAB60451.1; JOINED.
 DR EMBL; U14721; AAB60450.1; -;
 DR EMBL; U14720; AAB60450.1; JOINED.
 DR EMBL; U14721; AAB60448.1; -;
 DR EMBL; U13835; AAB60448.1; JOINED.
 DR EMBL; U14721; AAB60449.1; -;
 DR EMBL; U13835; AAB60449.1; JOINED.
 DR EMBL; X07539; CA30411.1; -;
 DR EMBL; X07539; CA30412.1; -;
 DR EMBL; X07540; CA30413.1; -;
 DR EMBL; X07540; CA30413.1; -;


```

DR MIM; 164690; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinaase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PRO0401; SH2DOMAIN.
DR PRINTS; PRO0452; SH3DOMAIN.
DR PRINTS; PRO0109; TYRKINASE.
DR PRINTS; PRO00001; Prot_kinase; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50002; SH2; 1.
DR PROSITE; PS50001; SH3; 1.
DR TransErase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
KW Phosphorylation; SH2 domain; SH3 domain; Alternative splicing.
FT DOMAIN 107 167 SH2.
FT DOMAIN 173 263 SH3.
FT DOMAIN 288 539 PROTEIN KINASE.
FT DOMAIN 561 564 POLY-SER.
FT DOMAIN 658 660 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 732 739 POLY-GLY.
FT DOMAIN 843 1055 PRO-RICH.
FT DOMAIN 984 988 POLY-PRO.
FT NP BIND 294 302 ATP (BY SIMILARITY).
FT BINDING 317 317 ATP (BY SIMILARITY).
FT ACT_SITE 409 409 ATP (BY SIMILARITY).
FT MOD_RES 439 439 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT VARSPDIC 1 73 MGOQVRGVEAPGLQOPQOPGIRGSSAARSGRRDPAGT
TERTGFIPTQHDHPASCVEDEGFGDRTGSSP -> MVLGT
VLLPNTYGRDDQTSCLCCTEASESALPDLT (in
isoform 1A).
SQ SEQUENCE 1182 AA; 128343 MW; ED93869BC2B14FPA CRC64;
Query Match 28.1%; Score 365.5; DB 1; Length 1182;
Best Local Similarity 35.0%; Pred. No. 4.2e-22;
Matches 89; Conservative 47; Mismatches 101; Indels 17; Gaps 9;
QY 1 YRAKWIISQDEKVAVKL---LKIE--KEAEILSVLSHRIIOFYGV-ILPEPNYGIIVT 52
DB 303 YGVGVKKYSITVAVKTLKEDTMEVEFLKEAAVKEIKHNLVQLGVCLEPPFY-IVT 361
QY 53 EYASLSGLVDYINSNSEKMDHIMTATVDKAKGHYLMHAPVKYIHRDLKSRNVIA 112
DB 362 EYMPYGNLADYLRKCNREEVTAVALYMAQISSAMEYLEKK--NFIHRDLAARCLVG 418
QY 113 ADGVLCICDGRGAFNHTHMSLVGT-PP--MMAPGVLOSLEVSECTDYSGVILMEN 169
DB 419 ENNVAVKADGSLRLMTGDTYTHAAGKFKITWTAPSLAVNTFSIKSDVMAFGVILMET 478
QY 170 LTRFV-PFKGLEGLQVAMLVENKNERLTISSCFSEFALLHOCWEADAKKRSFKQIIS 228
DB 479 ATYGMSPPYPIIDISQV-YDLEKGYMNEQPEGCPKYVIELMRACWKSPPADRSPFAETHQ 537
QY 229 ILESMSNDTSLPDK 242
DB 538 AFETMFHDSISHE 551
RESULT 13
ABL_DROME STANDARD; PRT; 1520 AA.
AC P00522;
DT 21-JUL-1986 (Rel. 01, Created)

```

```

DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DR Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-ash).
GN ABL OR DASH OR ABL-1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88174728; PubMed=2832740;
RA Henkemeyer M.J., Bennett R.L., Gertler F.B., Hoffmann F.M.;
RT "DNA sequence, structure, and tyrosine kinase activity of the
RT Drosophila melanogaster Abelson proto-oncogene homolog.";
RL Mol. Cell. Biol. 8:843-853(1988).
RN [2]
RP SEQUENCE OF 374-648 FROM N.A.
RX MEDLINE=84082064; PubMed=6317185;
RA Hoffmann F.M., Fresco L.D., Hoffmann F.M., Shilo B.-Z.;
RT "Nucleotide sequences of the Drosophila src and abl homologs:
RT conservation and variability in the src family oncogenes.";
RL Cell 35:393-401(1983).
RN [3]
RP FUNCTION.
RX MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Peifer M., Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
CC -1- FUNCTION: ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS
CC JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC this SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M19692; AAA28934.1; -.
DR EMBL; M19691; AAA28934.1; JOINED.
DR EMBL; M19691; AAA28934.1; JOINED.
DR EMBL; K01042; AAA28443.1; -.
DR PIR; A28128; TVFPA.
DR HSSP; P00519; IAB2.
DR FlyBase; FBgn0000017; Abl.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0018977; C:extrinsic to plasma membrane; IDA.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR GO; GO:0005927; C:muscle tendon junction; IDA.
DR GO; GO:0004713; F:protein tyrosine kinase activity; IDA.
DR GO; GO:0007411; P:axon guidance; IMP.
DR GO; GO:0007417; P:central nervous system development; IGI.
DR GO; GO:0007391; P:dorsal closure; NAS.
DR GO; GO:0002009; P:morphogenesis of an epithelium; NAS.
DR GO; GO:0006360; P:regulation of cell shape; NAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinaase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PRO0401; SH2DOMAIN.

```

```

DR PRINTS; PRO0452; SH3DOMAIN.
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW SH2 domain; SH3 domain.
KW DOMAIN 204 265 SH3.
FT DOMAIN 271 363 SH2.
FT DOMAIN 388 644 PROTEIN_KINASE.
FT NP_BIND 394 402 ATP (BY SIMILARITY).
FT BINDING 417 417 ATP (BY SIMILARITY).
FT ACT_SITE 509 509 BY SIMILARITY.
FT MOD_RES 539 539 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 374 377 LSPE -> ASAQ (IN REF. 2).
FT CONFLICT 645 648 ESS1 -> VGDV (IN REF. 2).
SQ SEQUENCE 1520 AA; 161836 MW; AD6A5060579FAD7B CRC64;

Query Match 28.0%; Score 364.5; DB 1; Length 1520;
Best Local Similarity 34.1%; Pred. No. 6.8e-22;
Matches 86; Conservative 47; Mismatches 104; Indels 15; Gaps 7;

QY 1 YRAKWTISQDEKVEVAVK-----LK-IEKEAEILSVSHNIIQFGVILEPNNYGIATVE 53
DB 403 YEVAVKRYGVTAVKTKEDTKMALKDFLEAAIMKEKKNVOLLGVCTREPPFIITTE 462
QY 54 YASLGSLYDYINSNRSEMDMHIWATDVAKGMHYLHMAEPVKYIHRDLKSRNVVIA 113
DB 463 FMHGNLDFLRSGRETLVAALLVVAQTIGASMSYLSER---NVIHRDLARNCLVGD 519
QY 114 DGVLTICDFASRPHNHTTMSLVGT-FP--WMAPEYIGSLPYSEICDYSYGVVLMEML 170
DB 520 NKLIVKVADEFLARMDRTYTAHAGAFPIKMTAPBEELAVNKRSTSDVWAFGLWLEIA 579
QY 171 TRFV-PEKLEGQVAMLVYKNEERLIPSSCRFAELHOCMEADAKRPSFKQIISI 229
DB 580 TYGMSPIPAID-LTDVYHKLDKGYRMRPCCPEPEYVDLMKQCWMDATDRPFKSIHNA 638
QY 230 LESMSNDTSLPD 241
DB 639 LEHMFQESSITE 650

RESULT 14
KYKI_DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
protein kinase 1).
DE PYKA OR SPLA OR DPYKL.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10.
RX MEDLINE=97053827; PubMed=8898241;
RA Nickols G.H., Oshrov N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase spla is essential for
spore differentiation.";
RL Development 122:3295-3305 (1996).
RN [2]

```

```

RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583 (1990).
CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32174; AAB41125.1; -.
DR EMBL; M33785; AAA33202.1; -.
DR PIR; T18276; T18276.
DR DICTYDB; DD03010; PYKA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR003877; SPRY_receptor.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00449; SPRY; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 460 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 532 532 POLY-ASN.
FT DOMAIN 596 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
FT DOMAIN 1195 1210 POLY-ASN.
FT DOMAIN 1215 1220 POLY-GLN.
FT DOMAIN 1224 1233 POLY-GLN.
FT DOMAIN 1266 1274 POLY-PRO.
FT DOMAIN 1289 1561 PROTEIN_KINASE.
FT NE_BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT_SITE 1417 1417 D -> R (IN REF. 2).
FT CONFLICT 1248 1248 V -> L (IN REF. 2).
FT CONFLICT 1435 1435 V -> L (IN REF. 2).
SQ SEQUENCE 1584 AA; 174304 MW; SD1589458DBE01E3 CRC64;

Query Match 27.3%; Score 355.5; DB 1; Length 1584;
Best Local Similarity 34.1%; Pred. No. 3.9e-21;
Matches 93; Conservative 47; Mismatches 96; Indels 37; Gaps 8;

QY 2 RAKWISQDEKVEVAVK-----KLKIEKEAEILSVSHNIIQFGVILEP--PPVY 48
DB 1305 RGYW--RETDVAIKIYRPDQFTKSLVWFQNEVGILSLRHPNVVQFGACTAGGEDH 1362
QY 49 GIVTEYASLGSLYDYINSNRSEMDMHI-WTAWTDVAKGMHYLHMAEPVKYIHRDLKSR 107

```

```

Db 1363 C1VTEWGGGSLRQFLTDHNNLEQNPRIKLKALDIAGMNVLHGMTP-PIIHRLLSR 1421
QY 108 NVVI-----ADGVLKICDFGASRPHNHTTHM-SLVGTFPMAAEVYQSLP 152
Db 1422 NILDHNIDPKNPVSSRQDICKKISDFGISRLKKEQASQMTQSVCCIPMAEVEFKGDS 1481
QY 153 VSRCTQYSGVYLMVLTREVPFKGLEGLQVAMLVENKNERLTIPSSCFRFAELHOC 212
Db 1482 NSEKSPVYSTIGWVLFELLTSDPEQDMKPKMAMHLLAYESRPPILPTTSSKKEILTQC 1541
QY 213 WEADAKKRPSFKQIISLESMT-----SNDTSLP 240
Db 1542 WDSNPDSRPFQKIIIVHLKEMEDQGVSSFRASVP 1574

RESULT 15
ABL_FSVHY STANDARD; PRT; 439 AA.
ID ABL_FSVHY
AC P10447;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).
GN V-ABL.
OS Feline sarcoma virus (strain Hardy-Zuckerman 2).
OC Viruses; Retroviral viruses; Retroviridae; Gammaetrovirus.
OX NCBI_TaxID=11776;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87141338; PubMed=3029415;
RA Bergold P.J., Blumenthal J.A., D'Andrea E., Snyder H.W. Jr.,
RA Ledergerd L., Silverstone A., Nguyen H., Besmer P.;
RT "Nucleic acid sequence and oncogenic properties of the H22 feline
RT sarcoma virus v-abl insert.";
RL J. Virol. 61:1193-1202(1987).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL-POL
CC POLYPEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15805; AAA3042.1; -.
CC HSSP: P00519; 1BBZ.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; kinase_1.
DR Pfam: PF00017; SH2_1.
DR Pfam: PF00018; SH3_1.
DR PRINTS: PR00678; PI3KINASEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase_1.
DR PRODOM: PD000093; SH2_1.
DR PRODOM: PD000066; SH3_1.
DR SMART: SM00252; SH2_1.
DR SMART: SM00326; SH3_1.
DR SMART: SM00219; TyKc_1.

```

```

DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50011; SH2_1.
DR PROSITE: PS50001; SH2_1.
DR PROSITE: PS50002; SH3_1.
KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;
KW SH2 domain; SH3 domain.
FT DOMAIN 10 70. SH3.
FT DOMAIN 76 166. SH2.
FT DOMAIN 191 439. PROTEIN KINASE.
SQ SEQUENCE 439 AA; 50004 MW; 13579BDFED1481AB CRC64;

Query Match 27.0%; Score 351.5; DB 1; Length 439;
Best Local Similarity 32.9%; Pred. No. 1.8e-21;
Matches 78; Conservative 51; Mismatches 93; Indels 15; Gaps 7;

QY 1 YRAKMSQDEYAVVKL---LKIE---KEAEILSVLSHRNIIQFYGVILEPPNGIVTE 53
Db 206 YEGVWKKYSLTVAVTKLKEDTMEVEEFLKEAVMKIKHPNLVQLLGVCRTREPPFIITE 265
QY 54 YASLSGLVDYINSNRSEKMDHIMTWATDVAKGMYLHMEAPVKYIHRDLKSRNVVIA 113
Db 266 FMTYGNLIDYLRRCNRQEVNAVVLVMATQISSAMEYLEKK--NFIHRDLAARNCTVGE 322
QY 114 DGVTKICDRGASRPHNHTTHMSLVGT-PP--MMAEVEIOSLPVSECTQYSGVYLMVEMTL 170
Db 323 NHLVAVADGSLRLMTGDTYTAAGTFFPKTAPBSLAVNKFSLKSDVAVAGVLMEIA 382
QY 171 TREV-PFKGLEGLQVAMLVENKNERLTIPSSCFRFAELHOCWEADAKKRPSFKQI 226
Db 383 TYGMSYFPQIDLSQV-YELLEKDYMERPEGCEPKYELMRACWQNNPSDRAPFAEI 438

```

Search completed: December 5, 2003, 09:18:59
Job time : 8.03704 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:11:58 ; Search time 22.1667 Seconds
(without alignments)
2875.443 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277

Sequence: 1 YKAKISQDKVAVKKLKI.....SILESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeophages:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1300	100.0	455	4 Q9HCC4	Q9HCC4 homo sapien
2	1300	100.0	800	4 Q9NYE9	Q9NYE9 homo sapien
3	1300	100.0	800	4 Q9NYL2	Q9NYL2 homo sapien
4	1300	100.0	800	4 Q9HCC5	Q9HCC5 homo sapien
5	1300	100.0	800	4 Q9HDD2	Q9HDD2 homo sapien
6	1289	99.2	454	11 Q9ESL3	Q9ESL3 mus musculu
7	1289	99.2	802	11 Q9ESL4	Q9ESL4 mus musculu
8	1281	98.5	289	11 Q9BR73	Q9BR73 mus musculu
9	1064	81.8	371	13 Q9OZY8	Q9OZY8 brachydactylo
10	511.5	39.3	406	10 Q23719	Q23719 arabidopsis
11	511.5	39.3	880	10 Q8S9K4	Q8S9K4 arabidopsis
12	508.5	38.7	1066	4 Q9H2N5	Q9H2N5 homo sapien
13	503.5	38.0	608	11 Q8B1G8	Q8B1G8 mus musculu
14	494	38.0	888	11 Q8CDL6	Q8CDL6 mus musculu
15	494	38.0	1001	11 Q8VDG6	Q8VDG6 mus musculu
16	490	37.7	859	4 Q8WY25	Q8WY25 homo sapien

17	486	37.4	888	11 Q8CBX3	Q8CBX3 mus musculu
18	483	37.2	564	4 Q9H1Y7	Q9H1Y7 homo sapien
19	482.5	37.1	847	4 Q16584	Q16584 homo sapien
20	481.5	37.0	850	11 Q9J0J5	Q9J0J5 mus musculu
21	480	36.9	886	10 Q9LYI8	Q9LYI8 arabidopsis
22	475.5	36.6	1030	10 Q8L6Z5	Q8L6Z5 arabidopsis
23	475.5	36.6	1030	10 Q9C9U5	Q9C9U5 arabidopsis
24	474	36.5	570	4 Q8WMN2	Q8WMN2 homo sapien
25	474	36.5	1036	4 Q8WMN1	Q8WMN1 homo sapien
26	472.5	36.3	977	4 Q9WY24	Q9WY24 drosophila
27	469.5	36.1	966	4 Q43283	Q43283 homo sapien
28	460.5	35.4	462	10 Q39886	Q39886 glycine max
29	453	34.8	982	10 Q65833	Q65833 lycopersico
30	452	34.8	903	10 Q9FPR5	Q9FPR5 oryza sativ
31	448	34.5	412	10 Q9M0S5	Q9M0S5 arabidopsis
32	445.5	34.3	1148	5 Q9SVF6	Q9SVF6 drosophila
33	445.5	34.3	1161	5 Q95UN8	Q95UN8 drosophila
34	444	34.2	992	10 Q8LPH3	Q8LPH3 arabidopsis
35	441	33.9	525	10 Q9F1I6	Q9F1I6 arabidopsis
36	440	33.8	982	10 Q93YG8	Q93YG8 lycopersico
37	438	33.7	525	10 Q9M0Z2	Q9M0Z2 arabidopsis
38	436.5	33.6	933	10 Q9FPR3	Q9FPR3 arabidopsis
39	436.5	33.6	1161	5 Q8WRK7	Q8WRK7 drosophila
40	436	33.5	417	10 Q8GVJ0	Q8GVJ0 oryza sativ
41	435.5	33.5	847	10 Q93XL9	Q93XL9 rosa hybrid
42	434	33.4	855	5 Q01700	Q01700 caenorhabdit
43	432	33.2	957	10 Q9FPR4	Q9FPR4 hordeum vul
44	431.5	33.2	553	10 Q81808	Q81808 arabidopsis
45	431.5	33.2	570	10 Q8RWL6	Q8RWL6 arabidopsis

ALIGNMENTS

RESULT 1

ID Q9HCC4 PRELIMINARY; PRT; 455 AA.

AC Q9HCC4; Q9HCC4; 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper containing kinase AZK) (Mixed lineage kinase) (Mixed lineage kinase-related kinase MKK-beta).

GN MLTK.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Gotoh I., Adachi M., Nishida E.;

RT "Identification and Characterization of a Novel MAP Kinase Kinase Kinase, MLTK.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RA Strausberg R.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Acton S.;

RT "MLK-mixed lineage kinase.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=21950776; PubMed=11936244;

RA Gross E.A., Callow W.G., Waldbaum U., Thomas S., Ruggieri R.;

RT "WRK, a Mixed Lineage Kinase-Related Molecule That Plays a Role in gamma-Radiation-Induced Cell Cycle Arrest.";

RL J. Biol. Chem. 277:13873-13882(2002).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AB049734; BAB16445.1; -
 DR EMBL: BC001401; AA010401.1; -
 DR EMBL: AF328454; AK1615.1; -
 DR EMBL: AF480462; AAL85892.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 455 AA; 51582 MW; E87DB844D58B752 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 455;
 Best Local Similarity 100.0%; Pred. No. 1.3e-114;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMSQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60
 DB 31 YRAKMSQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90
 QY 61 YDINSNRSEMDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
 DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
 QY 121 DFGASRFHNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPKGL 180
 DB 151 DFGASRFHNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPKGL 210
 QY 181 GLQVAMLVVEKNERLTIPTSCPRSPFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 240
 DB 211 GLQVAMLVVEKNERLTIPTSCPRSPFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 270
 QY 241 DKCNSFL 247
 DB 271 DKCNSFL 277

RESULT 2
 ID Q9NYE9 PRELIMINARY; PRT; 800 AA.
 AC Q9NYE9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed lineage kinase-related kinase MKK-alpha).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McNeen J.J., Fritma N., Diamond T.E., Dower S.K., Guesdon F.;
 RT "Cloning and characterisation of AZK, a mixed lineage kinase containing a sterile-alpha motif."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21950776; PubMed=11836244;
 RX Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;
 RT "MRK, a Mixed lineage kinase-related Molecule That Plays a Role in gamma-Radiation-induced Cell Cycle Arrest."
 RT J. Biol. Chem. 277:13873-13882(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF251441; AAF65822.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR001660; SAM.

DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00454; SAM; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50105; SAM_DOMAIN; 1.
 KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91181 MW; B289D836EC52E295 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;
 Best Local Similarity 100.0%; Pred. No. 2.7e-114;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMSQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60
 DB 31 YRAKMSQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90
 QY 61 YDINSNRSEMDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
 DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
 QY 121 DFGASRFHNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPKGL 180
 DB 151 DFGASRFHNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPKGL 210
 QY 181 GLQVAMLVVEKNERLTIPTSCPRSPFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 240
 DB 211 GLQVAMLVVEKNERLTIPTSCPRSPFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 270
 QY 241 DKCNSFL 247
 DB 271 DKCNSFL 277

RESULT 3
 ID Q9NYL2 PRELIMINARY; PRT; 800 AA.
 AC Q9NYL2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Mixed lineage kinase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20384179; PubMed=10924358;
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,
 RA Chou C.K., Yang J.J.;
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein containing a leucine-zipper and a sterile-alpha motif."
 RL Biochem. Biophys. Res. Commun. 274:811-816(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF238255; AAF63490.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00454; SAM; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50105; SAM_DOMAIN; 1.

KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 91264 MW; DA82D7AB82082F43 CRC64;
Query Match 100.0%; Score 1300; DB 4; Length 800;
Best Local Similarity 100.0%; Pred. No. 2.7e-114;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 60
DB 31 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 90
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
QY 121 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 180
DB 151 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 210
QY 181 GLQVAMLVVEKNERLTIPSSCFPSFAELLHQCEADAKRPSFKQIISLESMSNDTSLP 240
DB 211 GLQVAMLVVEKNERLTIPSSCFPSFAELLHQCEADAKRPSFKQIISLESMSNDTSLP 270
QY 241 DKCNSTFL 247
DB 271 DKCNSTFL 277
RESULT 4
Q9HCC5 PRELIMINARY; PRT; 800 AA.
ID Q9HCC5
AC Q9HCC5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MLTK-alpha.
GN MLTK.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; PubMed=11042189;
RA Gotoh I., Adachi M., Nishida E.,
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049733; BAB16444.1; -.
DR HSSP; P12931; IFMK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FAD0 CRC64;
Query Match 100.0%; Score 1300; DB 4; Length 800;
Best Local Similarity 100.0%; Pred. No. 2.7e-114;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 60
DB 31 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 90

QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
QY 121 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 180
DB 151 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 210
QY 181 GLQVAMLVVEKNERLTIPSSCFPSFAELLHQCEADAKRPSFKQIISLESMSNDTSLP 240
DB 211 GLQVAMLVVEKNERLTIPSSCFPSFAELLHQCEADAKRPSFKQIISLESMSNDTSLP 270
QY 241 DKCNSTFL 247
DB 271 DKCNSTFL 277
RESULT 5
Q9HDD2 PRELIMINARY; PRT; 800 AA.
ID Q9HDD2
AC Q9HDD2
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Placuble mixed-lineage kinase protein.
GN MLTKAK.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid organ;
RA Abe Y., Ueda N.;
RT "Placible Mixed-lineage kinase derived from LAK cell.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB030034; BAB12040.1; -.
DR HSSP; P12931; IFMK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 91155 MW; B2814509EC54B07A CRC64;
Query Match 100.0%; Score 1300; DB 4; Length 800;
Best Local Similarity 100.0%; Pred. No. 2.7e-114;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 60
DB 31 YRAKWIISQDEKAVVKKLLKTEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEYASLGS 90
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
QY 121 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 180
DB 151 DFGASRFHNHTTMSLVGTFPPMAAPEVIOSLPVSECTDYTSYGVVLMEMLTREVPFKGLE 210
QY 181 GLQVAMLVVEKNERLTIPSSCFPSFAELLHQCEADAKRPSFKQIISLESMSNDTSLP 240

```

Db      211 GLQVAMLVVEKNERLTI PSSCPSPFAELHQCWEADAKKRPSFKQIISLESMSNDTSLP 270
QY      241 DKCNSFL 247
Db      271 DKCNSFL 277

RESULT 6
Q9ESL3
ID      09ESL3      PRELIMINARY;      PRT;      454 AA.
AC      09ESL3;
DT      01-MAR-2001 (TReMBLrel. 16, Created)
DT      01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      MLTK-beta.
GN      ZAK OR MLTK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=2164927; PubMed=11042189;
RA      Gotch I., Adachi M., Nishida E.;
RT      "Identification and Characterization of a Novel MAP Kinase Kinase
RT      Kinase, MLTK."
RL      J. Biol. Chem. 276:4276-4286(2001).
CC      -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR      EMBL; AB049731; BAB16443.1; -.
DR      HSSP; P12931; IFMK.
DR      MGD; MGI:1931274; Zak.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR001245; Tyr_pkinase.
DR      Pfam; PF00069; pkinase; 1.
DR      PRINTS; PR00109; TYRKINASE.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ      SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

Query Match      99.2%; Score 1289; DB 11; Length 454;
Best Local Similarity 98.8%; Pred. No. 1.4e-113;
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 YRAKWTISQKEVAVKKLKIEKEAELISVLSHRNIIQFVGIVLEPPNYGIVTEYASLSGL 60
Db      31 YRAKWTISQKEVAVKKLKIEKEAELISVLSHRNIIQFVGIVLEPPNYGIVTEYASLSGL 90
QY      61 YDIYINSRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
Db      91 YDIYINSRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
QY      121 DFGASRFHNHTTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180
Db      151 DFGASRFHNHTTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210
QY      181 GLQVAMLVVEKNERLTI PSSCPSPFAELHQCWEADAKKRPSFKQIISLESMSNDTSLP 240
Db      211 GLQVAMLVVEKNERLTI PSSCPSPFAELHQCWEADAKKRPSFKQIISLESMSNDTSLP 270

QY      241 DKCNSFL 247
Db      271 DKCNSFL 277

RESULT 7
Q9ESL4
ID      09ESL4      PRELIMINARY;      PRT;      802 AA.
AC      09ESL4;
DT      01-MAR-2001 (TReMBLrel. 16, Created)
DT      01-MAR-2001 (TReMBLrel. 16, Last sequence update)

```

```

DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      MLTK alpha (Sterile-alpha motif and leucine zipper containing kinase
DE      AZK).
GN      ZAK OR MLTK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=2164927; PubMed=11042189;
RA      Gotch I., Adachi M., Nishida E.;
RT      "Identification and Characterization of a Novel MAP Kinase Kinase
RT      Kinase, MLTK."
RL      J. Biol. Chem. 276:4276-4286(2001).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N;
RA      Strausberg R.;
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR      EMBL; AB049731; BAB16442.1; -.
DR      EMBL; BC023718; AAH23718.1; -.
DR      HSSP; P12931; IFMK.
DR      MGD; MGI:1931274; Zak.
DR      InterPro; IPR000194; ATPase_a/bcentre.
DR      InterPro; IPR00719; Prot_kinase.
DR      InterPro; IPR001660; SAM.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      Pfam; PF00069; pkinase; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00454; SAM; 1.
DR      PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      PROSITE; PS00105; SAM_DOMAIN; 1.
KW      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ      SEQUENCE 802 AA; 91719 MW; D431DF8F312A43CC CRC64;

Query Match      99.2%; Score 1289; DB 11; Length 802;
Best Local Similarity 98.8%; Pred. No. 3e-113;
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 YRAKWTISQKEVAVKKLKIEKEAELISVLSHRNIIQFVGIVLEPPNYGIVTEYASLSGL 60
Db      31 YRAKWTISQKEVAVKKLKIEKEAELISVLSHRNIIQFVGIVLEPPNYGIVTEYASLSGL 90
QY      61 YDIYINSRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
Db      91 YDIYINSRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
QY      121 DFGASRFHNHTTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180
Db      151 DFGASRFHNHTTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210
QY      181 GLQVAMLVVEKNERLTI PSSCPSPFAELHQCWEADAKKRPSFKQIISLESMSNDTSLP 240
Db      211 GLQVAMLVVEKNERLTI PSSCPSPFAELHQCWEADAKKRPSFKQIISLESMSNDTSLP 270

QY      241 DKCNSFL 247
Db      271 DKCNSFL 277

RESULT 8
Q9BR73
ID      09BR73      PRELIMINARY;      PRT;      289 AA.
AC      09BR73;
DT      01-MAR-2003 (TReMBLrel. 23, Created)
DT      01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Sterile-alpha motif and leucine zipper containing kinase AZK.

```


OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue=Brain;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK045444; BAC32371.1; -
 SQ SEQUENCE 289 AA; 32872 MW; 550B5651066A0463 CRC64;

Query Match 98.5%; Score 1281; DB 11; Length 289;
 Best Local Similarity 98.4%; Pred. No. 4.5e-113;
 Matches 243; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGIETVYASLGL 60
 DB 31 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGIETVYASLGL 90
 QY 61 YDYINSRSEEMDMHMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIADGVLTIC 120
 DB 91 YDYINSRSEEMDMHMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIADGVLTIC 150
 QY 121 DFGASRFHNHTTMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLEMLTREVPKGL 180
 DB 151 DFGASRFHNHTTMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLEMLTREVPKGL 210
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEWADAKKRPSPKQIISLESMSNDTSLP 240
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEWADAKKRPSPKQIISLESMSNDTSLP 270

QY 241 DKCNSTL 247
 DB 271 DQCNSTL 277

RESULT 9
 Q90ZY8 PRELIMINARY; PRT; 371 AA.
 AC Q90ZY8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Protein kinase Npk.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NX NCBI_TaxID=7955;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Chou C.-M., Lee I.-L., Lau J.-H., Huang C.-J.;
 RT "A novel protein kinase, zNPK, from the zebrafish."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF265343; AK52416.1; -
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 371 AA; 42456 MW; 9B918B8A8B20D96 CRC64;

Query Match 81.8%; Score 1064; DB 13; Length 371;
 Best Local Similarity 79.8%; Pred. No. 2.2e-92;
 Matches 197; Conservative 23; Mismatches 27; Indels 0; Gaps 0;

QY 1 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGIETVYASLGL 60
 DB 62 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGIETVYASLGL 121
 QY 61 YDYINSRSEEMDMHMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIADGVLTIC 120
 DB 122 YELSSADSEEMDMQVMTWAMEIKGMHYLMEAPVKYIHRDLKSRNVVIADGVLTIC 181
 QY 121 DFGASRFHNHTTMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLEMLTREVPKGL 180
 DB 182 DFGASRFHNHTTMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLEMLTREVPKGL 241
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEWADAKKRPSPKQIISLESMSNDTSLP 240
 DB 242 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEWADAKKRPSPKQIISLESMSNDTSLP 301
 QY 241 DKCNSTL 247
 DB 302 DQCNSTL 308

RESULT 10
 O23719 PRELIMINARY; PRT; 406 AA.
 AC O23719;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE MAPK delta-1 protein kinase (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 NX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=99196996; PubMed=10095117;
 RA Jounanvic S., Hamal A., Leprieux A.S., Tregear J.W., Kreis M.,
 RA Henry Y.;
 RT "Characterisation of novel plant genes encoding MEKK/STK1 and RAF-
 RT related protein kinases."
 RL Gene 229:171-181(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Jounanvic S., Leprieux A.S., Hamal A., Kreis M., Henry Y.;
 RT "Plant MAP kinase signalling pathways in the limekight."
 RL Adv. Bot. Res. 30:0-0(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: Y14199; CAA74591.1; -
 DR HSP: P12931; IFMK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON_TER
 SQ SEQUENCE 406 AA; 45965 MW; CE0D994BF2BC12AB CRC64;

Query Match 39.3%; Score 511.5; DB 10; Length 406;
 Best Local Similarity 40.5%; Pred. No. 5.3e-40;
 Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY 1 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGI 50
 DB 150 YRAKMTSQDEKAVVKKLLKIEKEAETLSVLSHRNIIQFGVILPEPNYGI 207

```
QY 51 VTEYASIGSLYDYINSRSEMDMHTMTATDVAKGMHYLHMEAPYKVIHRDLKSRNV 110
Db 208 LTFELPRGSLYRLH-RPNQOLDEKRRMVALVAKGMNYLHSHPT-VVHRDLKSPNLL 265
QY 111 IADGVLCICDGFASRFPHNT--THMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMW 168
Db 266 VDKMWVVKVCDPGLSRKHKHTYLSKSTACTPEMAPEVLARNPANEKCVYSGVILME 325
QY 169 MLTREVPFKLEGLOVAMLVENKNERLTTPSSCPSPFAELHOCWEADAKKPSFOIIS 228
Db 326 LATSRYPMKGLNPMQVAVGFQNRRLIEIPDDIDLTVAQIIRRCWQTEPHLRPSFTQLMQ 385
QY 229 ILESMSNDTSLPDKNS 245
Db 386 SLKRLQG-LNISNRANT 401

RESULT 11
Q8S9K4 PRELIMINARY; PRT; 880 AA.
AC Q8S9K4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE AT5G11850/F14F18.20.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach C.J., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Yoshizaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AY075641; AA177650.1;
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; pkinase.1.
DR ProDom: PD000001; Prot_kinase.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; transferase.
KW SEQUENCE 880 AA; 97881 MW; 5D9AD50C2B08444A CRC64;
SQ

Query Match 39.3%; Score 511.5; DB 10; Length 880;
Best Local Similarity 40.5%; Pred. No. 1,5e-39;
Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY 1 YRAKWSIQDEKVAVKL-----KIEKEAELISVLSHRNIIOFYGVILEPPNYGI 50
Db 624 YRAEW--NGTEVAVKFLDDPFGDALTPKSELEIMLRHNVILFKAAVTRPREFSI 681
QY 51 VTEYASIGSLYDYINSRSEMDMHTMTATDVAKGMHYLHMEAPYKVIHRDLKSRNV 110
Db 682 LTFELPRGSLYRLH-RPNQOLDEKRRMVALVAKGMNYLHSHPT-VVHRDLKSPNLL 729
QY 111 IADGVLCICDGFASRFPHNT--THMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMW 168
Db 740 VDKMWVVKVCDPGLSRKHKHTYLSKSTACTPEMAPEVLARNPANEKCVYSGVILME 799
QY 169 MLTREVPFKLEGLOVAMLVENKNERLTTPSSCPSPFAELHOCWEADAKKPSFOIIS 228
Db 386 SLKRLQG-LNISNRANT 401
```

```
Db 800 LATSRYPMKGLNPMQVAVGFQNRRLIEIPDDIDLTVAQIIRRCWQTEPHLRPSFTQLMQ 859
QY 229 ILESMSNDTSLPDKNS 245
Db 860 SLKRLQG-LNISNRANT 875

RESULT 12
Q9H2N5 PRELIMINARY; PRT; 1066 AA.
AC Q9H2N5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McNeen J.C., Dower S.K., Guesdon F.;
RT "cDNA sequence and gene organisation of mixed lineage kinase 1."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL: AF251442; AKG44591.1; -.
DR HSSP: P29355; ISEM.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00069; pkinase.1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase.1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00002; SH3; 1.
KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW transferase.
FT NON TER
SQ SEQUENCE 1066 AA; 118463 MW; EDD08BEEB7482723 CRC64;

Query Match 39.1%; Score 508.5; DB 4; Length 1066;
Best Local Similarity 42.7%; Pred. No. 3,6e-39;
Matches 108; Conservative 43; Mismatches 77; Indels 25; Gaps 4;

QY 1 YRAKWSIQDEKVAVKL-----KILKIEKEAELISVLSHRNIIOFYGVILEPPNY 48
Db 107 YRAFWIGD--EVAVNAARHPDDEDISQTEENRQEKALAMLKHPRIILMRGVCLEKPEVL 164
QY 49 GIVTEYASIGSLYDYINSRSEMDMHTMTATDVAKGMHYLHMEAPYKVIHRDLKSRNV 108
Db 165 CLVMEFARAGPILNRVLSGR--IPDDILVNMAVOIARGMNYLHDEAIVPIIHRDLKSSN 221
QY 109 VVI-----AADGVLCICDGFASRFPHNTTHMSLVGTFPMMAPEVIOQLPVSETCDTY 160
Db 222 IILQKVENGDLSNKLTKITDFGLAEWHRTTKMSAAGYVAMVAPVIRASMPKSGSDVW 281
QY 161 SYGVVLMELTREVPFKLEGLOVAMLVENKNERLTTPSSCPSPFAELHOCWEADAKK 220
Db 282 SYGVVLMELLTGEVPRFGIDGLAVAYGVAMNKLALPIPTCEBPFAKLMEDCMNPDRSR 341
QY 221 PSFQIISILESM 233
Db 342 PSFTNILDQLTTI 354
```

```

RESULT 13
Q8BIG8      PRELIMINARY;      PRT;      608 AA.
AC 08BIG8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mixed lineage kinase MLK1 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK053843; BAC35552.1; -.
SQ SEQUENCE 608 AA; 68114 MW; 8619FCB47972E573 CRC64;

Query Match      38.7%; Score 503.5; DB 11; Length 608;
Best Local Similarity 42.3%; Pred. No. 5.1e-39;
Matches 107; Conservative 43; Mismatches 78; Indels 25; Gaps 4;

QY 1 YRAKWSIDKEVAV-----KLLKIEKAEIISVLSHRNIIOFYGVILEPPNY 48
DB 152 YRAFWAGD--EVAVKARHPDEDISKTENVAQEAFLFMLGHPNIALRGVCLKEPNL 209
QY 49 GYTFEYASIGSLVDYINSNSEMDMDHMTATDVAKGMHYLMKAPVKVYHRDLKSRN 108
DB 210 CLVMEFARGGFLNRVLGSKR---IPDILVMAVOIARGMNYLHDEAIVPIIHRDLKSSN 266
QY 109 VV-----AADGVLKICDFGASRFNHTHMSLVGTFPMMAPEVIQSLPSECTDY 160
DB 267 ILLQVENVGDDLSNKLTKITDPGLABEMHRTTMSAAGYAMAPVETIRASMPFSKSDW 326
QY 161 SYGVLMEMLTREVPFKGLEQLQVAMLVYKNERLTIPSSCPSPFAELLHQCEADAKKR 220
DB 327 SYGVLLMELLTGVRPGIDGLAVAGVANKKLALPIPTCPFPFAKLMDCNNPPPHSR 386
QY 221 PSFKQIISIESM 233
DB 387 PSFTSILDQLTTI 399

RESULT 14
Q8CDL6      PRELIMINARY;      PRT;      888 AA.
AC 08CDL6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mitogen activated protein kinase kinase kinase 12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK029882; BAC26658.1; -.
SQ SEQUENCE 888 AA; 96068 MW; 59D15FD840266383 CRC64;

```

```

Query Match      38.0%; Score 494; DB 11; Length 888;
Best Local Similarity 44.3%; Pred. No. 6.7e-38;
Matches 105; Conservative 44; Mismatches 78; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIOFYGVILEPPNYGYTFEYASIGSLVDYINSN 67
DB 180 EEVAVKAVKVDL-KETDIKHLRLKHPNITTFPGVCCQAGSYCIAMFCAGQGLYEVLRAG 238
QY 68 RSEEMDMHMTATDVAKGMHYLMKAPVKVYHRDLKSRNVIIAADGVLKICDFGASR- 126
DB 239 RPVTPESL--LVWMSMGIAAGMNYLH--KIHRDLKSPNMLITVDVVKISDFGTSKE 293
QY 127 FHNHTHMSLVGTFPMMAPEVIQSLPSECTDYSGYVLMEMLTREVPFKGLEQLQVAM 186
DB 294 LSDKSTWMSFAGVAMMAPEVIRNEPVSSEKVDIWSGVVLMELTGEIPIKVDVSAIIV 353
QY 187 LVYKNERLTIPSSCPSPFAELLHQCEADAKKRPFQIISIESMNDT-SLPDK 242
DB 354 GVSNSLHLPIVSSCPDGFKILLRCQMSKPRNRSFRQILHLHDIASADVLTSGE 410

RESULT 15
Q8VDG6      PRELIMINARY;      PRT;      1001 AA.
AC 08VDG6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 23, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase kinase 9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DB EMBL; BC021891; AA021891.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1001 AA; 109983 MW; E1042C868B953C CRC64;

Query Match      38.0%; Score 494; DB 11; Length 1001;
Best Local Similarity 39.4%; Pred. No. 7.8e-38;
Matches 108; Conservative 47; Mismatches 85; Indels 34; Gaps 5;

QY 1 YRAKWSIDKEVAVKLLK-----IEKEAEIISVLSHRNIIOFYGVILEPPNY 48
DB 125 YRATW--QCQEVAVKARDPEDODAAAASVAREARLFLAMLRHPNIIOLRGVCLQPHL 182
QY 49 GYTFEYASIGSLVDY-----NSNRSEMDMDHMTATDVAKGMHYLMKAPVK 98
DB 183 CLVLEFARGGALRLAALAAADPRAPGPRRARIPOVIVMAVQIARGLYLHEEAVVP 242
QY 99 VIHRDLKSRNVV-----AADGVLKICDFGASRFNHTHMSLVGTFPMMAPEVIQ 150
DB 243 IHRDLKSSNILLKIEIHDDICNKLTKITDPGLABEMHRTTMSAAGYAMMAPEVINS 302

```

Qy	151	LPVSEICDITYSGVWLMEMLTREVPFKGLEQVAMLVKMERLITPSCPRFALLH	210
Db	303	SIFSXSDDIWSYGVLMWELLTGEVPRGIDGLAVAGVAVNKLTLPIPTCPEPFAKLWK	362
Qy	211	OCWEADAKKRPSPFKOIIISLESMSND--TSLPDK	242
Db	363	ECWEQDPHIRPSPALLIQLTALBEAVLTNMQE	396

Search completed: December 5, 2003, 09:20:16
 Job time : 23.1667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 08:30:40 ; Search time 26.7407 Seconds
(without alignments)
1466.132 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277
Perfect score: 1300
Sequence: 1 YKAKMSQKQEVAVKKLTKI.....SILESMSNDTSLPDKCNSTL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2004.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1300	100.0	455	21	AA18657
2	1300	100.0	455	21	AA18657
3	1300	100.0	455	21	AA18657
4	1300	100.0	473	22	AA18657
5	1300	100.0	800	22	AA18657
6	1300	100.0	800	22	AA18657
7	1300	100.0	800	22	AA18657
8	1253	96.4	349	22	AA18657
9	744	57.2	141	21	AA18657

10	510.5	39.3	394	24	AA18657	Human protein kinase
11	509.5	39.2	1024	22	AA18657	Human NOV7, a mixe
12	497.5	38.3	1046	22	AA18657	Human kinase (PKIN
13	497.5	38.3	1097	23	AA18657	Human PKIN-12 prot
14	490	37.7	859	18	AA18657	Human leucine-zipp
15	490	37.7	859	18	AA18657	Human leucine-zipp
16	490	37.7	859	24	AA18657	Polypeptide sequen
17	489	37.6	888	23	AA18657	Mouse ischaemic co
18	487.5	37.5	1021	23	AA18657	Novel human protei
19	483	37.2	719	22	AA18657	Human protein kinase
20	483	37.2	1036	24	AA18657	Novel human protei
21	483	37.2	1036	24	AA18657	Amino acid sequen
22	483	37.2	1036	24	AA18657	Human kinase and p
23	482.5	37.1	847	23	AA18657	Human mitogen acti
24	472.5	36.3	977	22	AA18657	Drosophila melanog
25	470	36.2	144	22	AA18657	Novel human diago
26	469.5	36.1	966	24	AA18657	Polypeptide sequen
27	469.5	36.1	1490	22	AA18657	Novel human diago
28	465	34.8	982	22	AA18657	Tomato TCFR2, lyc
29	452	34.8	903	22	AA18657	Rice EDRL Oryza
30	448.5	34.5	850	23	AA18657	Weicon constitutive
31	448	34.5	1369	21	AA18657	Arabidopsis thalia
32	448	34.5	374	21	AA18657	Arabidopsis thalia
33	448	34.5	412	21	AA18657	Arabidopsis thalia
34	445.5	34.3	1020	22	AA18657	Drosophila melanog
35	442	34.0	821	16	AA18657	Arabidopsis CTRL p
36	442	34.0	821	18	AA18657	Constitutive tripl
37	442	34.0	821	22	AA18657	Arabidopsis thalia
38	438	33.7	821	15	AA18657	Arabidopsis thalia
39	436.5	33.6	933	22	AA18657	Arabidopsis thalia
40	435.5	33.5	970	22	AA18657	Arabidopsis thalia
41	424	32.6	92	22	AA18657	Barley EDL1, Hord
42	424	32.6	92	22	AA18657	Novel central nerv
43	418.5	32.2	589	21	AA18657	Novel signal trans
44	418.5	32.2	732	21	AA18657	Arabidopsis thalia
45	418.5	32.2	760	21	AA18657	Arabidopsis thalia

ALIGNMENTS

RESULT 1	AA18657	standard; protein; 455 AA.
ID	AA18657	
AC	AA18657	
XX		
DT	22-JAN-2001	(first entry)
XX		
DE	A human regulator of intracellular phosphorylation.	
XX		
KW	Human; intracellular phosphorylation regulator; HRP; stroke; myeloma;	
KW	neurological disorder; Parkinson's disease; demyelinating disease;	
KW	meningitis; developmental disorder; neuromuscular disorder; cancer;	
KW	myasthenia gravis; cell proliferative disorder; actinic keratosis;	
KW	arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;	
KW	autoimmune disorder; inflammatory disorder; Addison's disease;	
KW	acquired immunodeficiency disease; allergy; diabetes mellitus;	
KW	rheumatoid arthritis; microbial infection; trauma.	
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	Domain	16..257
FT	Modified-site	/note="eukaryotic protein kinase domain"
FT	Modified-site	61
FT	Modified-site	/note="potential phosphorylation site"
FT	Modified-site	89
FT	Modified-site	/note="potential phosphorylation site"
FT	Modified-site	96
FT	Modified-site	/note="potential phosphorylation site"
FT	Modified-site	97
FT	Modified-site	/note="potential glycosylation site"

FT Binding-site 129..141
 FT /note= "protein kinase ATP-binding site"
 FT Modified-site 159
 FT /note= "potential glycosylation site"
 FT Modified-site 234
 FT /note= "potential phosphorylation site"
 FT Modified-site 252
 FT /note= "potential phosphorylation site"
 FT Modified-site 258
 FT /note= "potential phosphorylation site"
 FT Modified-site 265
 FT /note= "potential glycosylation site"
 FT Modified-site 268
 FT /note= "potential phosphorylation site"
 FT Region 294..322
 FT /note= "leucine zipper"
 FT Modified-site 302
 FT /note= "potential phosphorylation site"
 FT Modified-site 302
 FT /note= "potential phosphorylation site"
 FT Modified-site 342
 FT /note= "potential phosphorylation site"
 FT Modified-site 343
 FT /note= "potential phosphorylation site"
 FT Modified-site 346
 FT /note= "potential phosphorylation site"
 FT Modified-site 364
 FT /note= "potential phosphorylation site"
 FT Modified-site 409
 FT /note= "potential glycosylation site"
 FT Modified-site 410
 FT /note= "potential phosphorylation site"
 FT Modified-site 414
 FT /note= "potential phosphorylation site"
 FT Modified-site 415
 FT /note= "potential phosphorylation site"
 FT Modified-site 429
 FT /note= "potential phosphorylation site"
 FT Modified-site 434
 FT /note= "potential phosphorylation site"
 FT /note= "potential phosphorylation site"
 PN WO20005532-A2.
 XX
 XX
 PD 21-SEP-2000.
 XX
 XX 17-MAR-2000; 2000WO-US07277.
 XX
 XX 18-MAR-1999; 99US-0125593.
 PR 20-MAY-1999; 99US-0135049.
 PR 09-JUL-1999; 99US-0143188.
 XX
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DM, Au-Young J;
 PI
 XX WPI; 2000-602121/57.
 DR N-PSDB; AAA75674.
 XX
 XX Novel human intracellular phosphorylation regulator polypeptides and
 PT polynucleotides for diagnosis, prevention and treatment of
 PT neurological, cell proliferative and autoimmune/inflammatory disorders
 PT
 PT
 PT
 XX Claim 1; Page 75-76; 96pp; English.
 XX
 XX The present sequence represents a human regulator of intracellular
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and

CC other developmental disorders of the central nervous system,
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic Keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
 CC rheumatoid arthritis, microbial infection and trauma.
 XX
 XX Sequence 455 AA;
 SQ
 Query Match 100.0%; Score 1300; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 6,1e-128;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRAKWIISQDEKVAVKLLKIEKEAEILSVLSHRNIIOFYGVILLEPPNYGIVTEVASLSGL 60
 DB 31 YRAKWIISQDEKVAVKLLKIEKEAEILSVLSHRNIIOFYGVILLEPPNYGIVTEVASLSGL 90
 QY 61 YYTINSNRSEEDMDHMTWADVAKGMHYLMNEAPVYIHRDLKSRNVVIAADGVLTIC 120
 DB 91 YDINSNRSEEDMDHMTWADVAKGMHYLMNEAPVYIHRDLKSRNVVIAADGVLTIC 150
 QY 121 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVILMMLTREVPFKGLE 180
 DB 151 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVILMMLTREVPFKGLE 210
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEADAKKPSFKOIIISLESMSNDTSLP 240
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCEADAKKPSFKOIIISLESMSNDTSLP 270
 QY 241 DKCNSFL 247
 DB 271 DKCNSFL 277
 RESULT 2
 AA83278
 ID AA83278 standard; Protein; 455 AA.
 XX
 AC AA83278;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE Human survival regulating kinase (SRK).
 XX
 KW Survival regulating kinase; SRK; Raf; MAPK; MAPKK; MAPKK;
 KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
 KW apoptosis; cell survival; nuclear targeting; tumour; human;
 KW autoimmune disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200022142-A2.
 XX
 PD 20-APR-2000.
 XX
 XX 20-SEP-1999; 99WO-US22008.
 PF
 XX 13-OCT-1998; 98US-0104086.
 PR
 XX (ONVX-) ONVX PHARM INC.
 PA
 XX Ruggieri R, Callow M, Diaz P;
 PI
 XX WPI; 2000-317994/27.
 DR N-PSDB; AA93783.
 XX
 XX Novel human survival regulating kinase polypeptide for screening agents
 PT which modulate biological pathways associated with SRK useful in
 PT treating autoimmune diseases, tumors and apoptosis-related disorders
 XX
 XX Claim 4; Figure 2; 62pp; English.

XX Survival regulating kinases (SRK) are a class of proteins involved in
 CC cell signal transduction pathways such as mitogen-activated protein
 CC kinase pathways. A protein kinase activity means that the SRK can
 CC catalyse a reaction in which a phosphate group is transferred from a
 CC phosphate donor to a phosphate acceptor amino acid residue.
 CC preferably the hydroxyl side chain of a serine or threonine.
 CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
 CC activity is similar to that of a MAPKKK such as Raf. has a range of
 CC other activities including a cell growth-regulatory activity, a cell
 CC survival promoting activity, a HAX-1 binding activity, an apoptosis
 CC suppressing activity a MAPK activation or stimulatory activity, a
 CC nuclear targeting activity and a SRK-specific immunogenic activity.
 CC SRK is useful for identifying agents which modulate cellular
 CC transformations mediated by Ras and SRK and agents that modulate the
 CC apoptosis suppression activity of SRK. This information may be useful
 CC in the treatment of autoimmune diseases, tumours and apoptosis
 CC related disorders.

XX Sequence 455 AA;

Query Match 100.0%; Score 1300; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 6.1e-128;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWTISQKKEVAVKLLKTEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 60
 DB 31 YRAKWTISQKKEVAVKLLKTEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 90
 QY 61 YDINSNRSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120
 DB 91 YDINSNRSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 150
 QY 121 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
 DB 151 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210
 QY 181 GLOVAMLVKXKNERLTIPSSCPSPFAELHQCWEADAKRPSFKOITLESMSNDTSLP 240
 DB 211 GLOVAMLVKXKNERLTIPSSCPSPFAELHQCWEADAKRPSFKOITLESMSNDTSLP 270
 QY 241 DKCNSFL 247
 DB 271 DKCNSFL 277

RESULT 3
 ID AAY84321 standard; Protein; 455 AA.

XX AAY84321;
 XX
 XX 12-JUL-2000 (first entry)

DE A human cardiovascular system associated protein kinase-2.

XX Human; cardiovascular system associated protein kinase-2; CSAPK-2;
 KW signalling pathway; cell growth; cell differentiation; gene mapping;
 KM tissue typing; forensic identification; cardiovascular disease;
 XX congestive heart failure; transgenic animal.

OS Homo sapiens.

PN WO200014212-A1.

PD 16-MAR-2000.

PF 09-SEP-1999; 99WO-US20631.

XX 09-SEP-1998; 98US-0099657.

PR 29-SEP-1998; 98US-0163115.

PA (MILL-) MILLENNIUM PHARM INC.

XX Acton S;

DR WPI, 2000-271053/23.

DR N-PSDB; AA299726, AA299727.

PT New nucleic acid encoding cardiovascular system associated protein

PT kinase, used e.g. for diagnosis, treatment and prevention of

PT cardiovascular disease

PS Claim 2; Fig 2; 163pp; English.

XX The present sequence represents a human cardiovascular system associated
 CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling
 CC pathways associated with cell growth and differentiation. The CSAPK
 CC polypeptides and polynucleotides are used to screen for agents that
 CC specifically modulate CSAPK, which are potential therapeutic agents.
 CC They are also used for diagnosis, prognosis or monitoring of
 CC CSAPK-related diseases, gene mapping, tissue typing and forensic
 CC identification, and for treating or preventing disorders associated
 CC with aberrant CSAPK expression or activity, especially cardiovascular
 CC diseases such as congestive heart failure. They can also be used in
 CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate
 CC transgenic animals.

SO Sequence 455 AA;

Query Match 100.0%; Score 1300; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 6.1e-128;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWTISQKKEVAVKLLKTEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 60
 DB 31 YRAKWTISQKKEVAVKLLKTEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 90
 QY 61 YDINSNRSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120
 DB 91 YDINSNRSEEMDMHIMTWADVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 150
 QY 121 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
 DB 151 DFGASRFNHTHTMSLVGTFPMMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210
 QY 181 GLOVAMLVKXKNERLTIPSSCPSPFAELHQCWEADAKRPSFKOITLESMSNDTSLP 240
 DB 211 GLOVAMLVKXKNERLTIPSSCPSPFAELHQCWEADAKRPSFKOITLESMSNDTSLP 270
 QY 241 DKCNSFL 247
 DB 271 DKCNSFL 277

RESULT 4
 ID AAM25322 standard; Protein; 473 AA.

XX AAM25322;
 XX
 XX 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:837.

XX Human; cancer; HIV infection; human immunodeficiency virus;
 KW anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; virocid;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antilcer; osteoporotic; eczema;
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 OS Homo sapiens.
 PN MO200153455-A2.
 XX 26-UTL-2001.
 PD 22-DEC-2000; 2000MO-US35017.
 PF 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-457603/49.
 DR N-PSDB; AAB9263.
 XX Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 PT Claim 20; Page 191; 1217pp; English.
 XX AAB99166 to AAB99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: anti-inflammatory; antirheumatic;
 CC antirhectic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antitumagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
 CC antidiabetic; osteopathic; dermatological; antiallergic; antiaesthetic;
 CC antiparkinsonian; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX Sequence 473 AA;
 SQ
 Query Match 100.0%; Score 1300; DB 22; Length 473;
 Best Local Similarity 100.0%; Pred. No. 6, 5e-128;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRAKWSIQDEKVAVKLLIKKEAEILSVLSHRNIIQFVGVIIEPPNYGIVTEYASLGSU 60
 DB 49 YRAKWSIQDEKVAVKLLIKKEAEILSVLSHRNIIQFVGVIIEPPNYGIVTEYASLGSU 108
 QY 61 YDININRSEEMDMHIMTAVAGMHLHAEAPVKYVHRLKSRNVVIADGVLLKIC 120
 DB 109 YDININRSEEMDMHIMTAVAGMHLHAEAPVKYVHRLKSRNVVIADGVLLKIC 168
 QY 121 DFGASRFHNHTTMSLVGTFPPMMAPEVIOGLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180
 DB 169 DFGASRFHNHTTMSLVGTFPPMMAPEVIOGLPVSETCDTYSYGVVLMEMLTREVPFKGLE 228
 QY 181 GLQVAVLVVEKNERLITPSSCPSPFAELLHQCEADAKKRPSPKQIITILESNSNDTSLP 240
 DB 229 GLQVAVLVVEKNERLITPSSCPSPFAELLHQCEADAKKRPSPKQIITILESNSNDTSLP 288

QY 241 DKCNSFL 247
 DB 289 DKCNSFL 295
 RESULT 5
 ID AAB71957
 XX AAB71957 standard; Protein; 800 AA.
 AC AAB71957;
 XX 11-MAY-2001 (first entry)
 DT 11-MAY-2001 (first entry)
 DE Human TGF-beta receptor encoded by cDNA clone HDP5W48.
 XX Human; antisclerotic; dermatological; immunosuppressive; cytostatic;
 KW antiinflammatory; anti-HIV; immunostimulant; cardiant; vascular;
 KW ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;
 KW antiparkinsonian; antimicrobial; vulnerary; gene therapy; infection;
 KW transforming growth factor; TGF; TGF-beta receptor; immune disorder;
 KW hyperproliferative disorder; cardiovascular disease; angiogenesis;
 KW neurological disorder.
 XX Homo sapiens.
 OS
 XX MO200112670-A1.
 PN 22-FEB-2001.
 PD 10-AUG-2000; 2000MO-US21736.
 PF 13-AUG-1999; 99US-0148682.
 PR 20-SEP-1999; 99US-0154887.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Ni J;
 PI WPI; 2001-202858/20.
 DR N-PSDB; AAF75336.
 XX Nucleic acid molecules encoding 12 transforming growth factor-beta
 PT receptor polypeptides, useful for preventing, diagnosing and treating
 PT e.g. cancers, Parkinson's disease and diabetic retinopathy -
 PT Claim 11; Page 293-295; 311pp; English.
 XX The present sequence is one of 12 novel human transforming growth factor
 CC (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides
 CC and polypeptides may be used in the prevention, diagnosis and treatment
 CC of diseases associated with inappropriate polypeptide expression. Such
 CC diseases include immune disorders (e.g. multiple sclerosis, systemic
 CC lupus erythematosus and human immuno-deficiency virus (HIV) infections),
 CC hyperproliferative disorders (e.g. cancers and Gaucher's disease),
 CC cardiovascular diseases (e.g. Scintlar syndrome, Chaga's cardiomyopathy
 CC and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)
 CC and infectious diseases. The polynucleotides and polypeptides are also
 CC useful for promoting wound healing, regeneration and/or chemotaxis. The
 CC polynucleotides and their complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples. The polypeptides may be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of protein expression and activity. The anti-TGF-beta receptor
 CC antibodies may be used to down regulate expression and activity and as
 CC diagnostic agents for detecting the presence of the polypeptides in
 CC samples.
 XX Sequence 800 AA;
 SQ
 Query Match 100.0%; Score 1300; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 1, 4e-127;

Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YAAKMTSQKKEVAVKLLKIEKAEILSVLSHRNIIQFVGVLLEPPNIGIVTEYASLGL 60
 DB 31 YAAKMTSQKKEVAVKLLKIEKAEILSVLSHRNIIQFVGVLLEPPNIGIVTEYASLGL 90
 QY 61 YDIINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIAADGVLKIC 120
 DB 91 YDIINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIAADGVLKIC 150
 QY 121 DFGASRFHNHTTMSLVGTFPPMMAPEVIOQLPVSETCDTYSYGVLWEMLTREVPKGL 180
 DB 151 DFGASRFHNHTTMSLVGTFPPMMAPEVIOQLPVSETCDTYSYGVLWEMLTREVPKGL 210
 QY 181 GIQVAMLVVEKNERLTIPTSCPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 240
 DB 211 GIQVAMLVVEKNERLTIPTSCPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 270
 QY 241 DKCNSTL 247
 DB 271 DKCNSTL 277

RESULT 6
 AAB65673
 ID AAB65673 standard; Protein; 800 AA.
 AC AAB65673;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 201.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiact; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN MO200073469-A2.
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000MO-US14842.
 XX
 PR 28-MAY-1999; 99US-0136503.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 DR WPI: 2001-032161/04.
 DB N-PSDB; AAF44701.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 PS Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 800 AA;

Query Match 100.0%; Score 1300; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 1.4e-127;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YAAKMTSQKKEVAVKLLKIEKAEILSVLSHRNIIQFVGVLLEPPNIGIVTEYASLGL 60
 DB 31 YAAKMTSQKKEVAVKLLKIEKAEILSVLSHRNIIQFVGVLLEPPNIGIVTEYASLGL 90
 QY 61 YDIINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIAADGVLKIC 120
 DB 91 YDIINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKYIHRDLKSRNVVIAADGVLKIC 150
 QY 121 DFGASRFHNHTTMSLVGTFPPMMAPEVIOQLPVSETCDTYSYGVLWEMLTREVPKGL 180
 DB 151 DFGASRFHNHTTMSLVGTFPPMMAPEVIOQLPVSETCDTYSYGVLWEMLTREVPKGL 210
 QY 181 GIQVAMLVVEKNERLTIPTSCPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 240
 DB 211 GIQVAMLVVEKNERLTIPTSCPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 270
 QY 241 DKCNSTL 247
 DB 271 DKCNSTL 277

RESULT 7
 ABP43736
 ID ABP43736 standard; Protein; 800 AA.
 AC ABP43736;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE Motif zipper containing kinase AZK.
 XX
 KW Neuroprotective; immunomodulator; cancer; chromosome 2q24.2;
 KW cytoskeletal; anti-inflammatory; gene therapy; nutritional supplement;
 KW wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KW vulnery.
 XX
 OS Homo sapiens.
 XX
 PN MO200231111-A2.
 PD 18-APR-2002.
 XX
 PF 11-OCT-2001; 2001MO-US27760.
 XX
 PR 12-OCT-2000; 2000US-0687527.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 DR WPI: 2002-426278/45.
 DB N-PSDB; AB060980.
 XX
 PT New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation -

XX Claim 20; SEQ ID # 639; 357bp + sequence listing; English.
PS
XX The invention relates to 446 newly isolated polynucleotide sequences.
CC The activity of polynucleotides of the invention may be described as,
CC vulnerability, neuroprotective, immunomodulator, cytostatic and
CC anti-inflammatory. Compositions comprising nucleic acids of the invention
CC are useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records
CC ABP43544-ABP43989 represent polypeptides encoded by polynucleotides of
CC the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 800 AA;
Query Match 100.0%; Score 1300; DB 23; Length 800;
Best Local Similarity 100.0%; Pred. No. 1.4e-127;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRAKMTISQDEKAVAKKLTKEKAEILSVLSHRNIIQFVGVLLEPPNYGIVTEYASLSGL 60
DB 31 YRAKMTISQDEKAVAKKLTKEKAEILSVLSHRNIIQFVGVLLEPPNYGIVTEYASLSGL 90
QY YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVYHRDLKSRNVVIAADGVLKIC 120
DB 91 YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVYHRDLKSRNVVIAADGVLKIC 150
QY 121 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDIYSYGVLMEMLTREVPFKGLE 180
DB 151 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDIYSYGVLMEMLTREVPFKGLE 210
QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPFKOIIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPFKOIIISLESMSNDTSLP 270
QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 8
AAG75571
ID AAG75571 standard; Protein; 349 AA.
AC AAG75571;
XX
XX 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6335.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA,
XX

XX WPI; 2001-235357/24.
DR N-PSDB; AAH34976.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 11; Page 7789-7790; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
SQ Sequence 349 AA;
Query Match 96.4%; Score 1253; DB 22; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.7e-123;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRAKMTISQDEKAVAKKLTKEKAEILSVLSHRNIIQFVGVLLEPPNYGIVTEYASLSGL 60
DB 87 YRAKMTISQDEKAVAKKLTKEKAEILSVLSHRNIIQFVGVLLEPPNYGIVTEYASLSGL 146
QY 61 YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVYHRDLKSRNVVIAADGVLKIC 120
DB 147 YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVYHRDLKSRNVVIAADGVLKIC 206
QY 121 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDIYSYGVLMEMLTREVPFKGLE 180
DB 207 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDIYSYGVLMEMLTREVPFKGLE 266
QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPFKOIIISLESMSNDTSL 239
DB 267 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPFKOIIISLESMSNDTSL 325

RESULT 9
AAG03583
ID AAG03583 standard; Protein; 141 AA.
AC AAG03583;
XX
XX 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 7664.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX

PR 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclet A, Giordano J;
XX
XX MPI; 2000-500381/45.
DR N-PSDB; AAC03589.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 13; SEQ ID 7664; 71bp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 141 AA;
XX
Query Match 57.2%; Score 744; DB 21; Length 141;
Best Local Similarity 98.6%; Pred. No. 3.9e-70;
Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 72 MDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNT 131
DB 1 MDMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRPHNT 60
XX
QY 132 THMSLVGTTPPMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLBGLQVAMLVEX 191
DB 61 THMSLVGTTPPMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLBGLQVAMLVEX 120
XX
QY 192 NERLTIPSSCPRSFAELHQC 212
DB 121 NERLTIPSSCPRSFAELHQC 141
XX
RESULT 10
AAG79648
ID AAG79648 standard; Protein; 394 AA.
XX
XX AAG79648;
AC
XX
DT 07-FEB-2003 (first entry)
XX
DE Human protein kinase-like protein identified from Swiss Prot.
XX
XX Human; protein kinase-like; chromosome 15; tyrosine kinase; enzyme;
KW class III; cancer; central nervous system; cardiovascular disorder;
KW chronic obstructive pulmonary disease; COPD; obesity; diabetes; CNS.
XX
XX Homo sapiens.
OS
XX
PN WO200281704-A2.
XX
PD 17-OCT-2002.
XX
PF 15-MAR-2002; 2002MO-EP02887.
XX
XX 16-MAR-2001; 2001US-276055P.
PR 24-SEP-2001; 2001US-324053P.
PR 03-OCT-2001; 2001US-326458P.

PR 10-DEC-2001; 2001US-337124P.
XX
XX (FARB) BAYER AG.
XX
XX Smolyar A;
PI
XX
XX MPI; 2003-040700/03.
DR
XX
XX New human protein kinase-like polypeptide for treating, preventing or
PT ameliorating cancer, central nervous system disorders, obesity,
PT diabetes, cardiovascular disorders and chronic obstructive pulmonary
PT disease -
XX
XX
PS Disclosure; Fig 3; 143pp; English.
XX
XX This sequence shows a human protein kinase-like polypeptide identified
CC from Swiss Prot. The protein kinase-like coding sequence is located on
CC chromosome 15. Related ESTs are expressed in lymphoma and germinal
CC centre B cells. The protein kinase-like protein has a likely function
CC as a tyrosine kinase class III protein. A reagent that modulates the
CC activity of protein kinase-like protein is useful in a medicament for
CC modulating the activity of human protein kinase-like protein in a
CC disease such as cancer, central nervous system (CNS) disorder, chronic
CC obstructive pulmonary disease (COPD), obesity, diabetes and
CC cardiovascular disorder. The reagent is also useful for treating a
CC human protein kinase-like dysfunction related disease including cancer,
CC CNS disorder, COPD, obesity, diabetes and cardiovascular disorder.
XX
SQ Sequence 394 AA;
XX
Query Match 39.3%; Score 510.5; DB 24; Length 394;
Best Local Similarity 42.7%; Pred. No. 7e-45;
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;
XX
QY 1 YRAKWSQDEKAVK-----KLLKTEKEAEILVLSHRNIIQFYGVILEPPNY 48
DB 18 YRAFWIGD--EVAVVAARHDPEDISQTEENVRQEKFLPAMLKHPHITLRGVCCKEPL 75
XX
QY 49 GIVTEYASLSGLSYDYINSRSEMDHMTWATDVAKGMHYLMEAPVKVIHRDLKSN 108
DB 76 CLMEFARAGPILNRVLSGR---IPEDILVNNAVOIARGMNYLHDEAIVPIIHRDLKSN 132
XX
QY 109 VVI-----AADGLKICDFGASRPHNTTHMSLVGFPPMAPPVIOQLPVSECDPY 160
DB 133 IILQKVENGDLSNKLTKTDFGLARBWRITTKAGAYAMAPVIVASPSKSDVA 192
XX
QY 161 SYGVLMEMLTREVPFKGLBGLQVAMLVEKNERLTIPSSCPRSFAELHQCMEADAKR 220
DB 193 SYGVLMEMLTREVPFKGLBGLQVAMLVEKNERLTIPSTCEPFAKLMEDCMNPPHNR 252
XX
QY 221 PSFKQIISTLES 233
DB 253 PSFTNILDQTLTI 265
XX
RESULT 11
ABB98408
ID ABB98408 standard; Protein; 1024 AA.
XX
XX ABB98408;
AC
XX
DT 21-OCT-2002 (first entry)
XX
XX Human NOV7, a mixed lineage kinase 2-like protein.
DE
XX
XX Human; NOV7; cytosolic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antischismatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes;
KW obesity; asthma; IGA nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;

KM wasting disorder; mixed lineage kinase 2-like protein.
 XX Homo sapiens.
 OS
 PN WO200255704-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US00554.
 XX
 PR 09-JAN-2001; 2001US-260417P.
 PR 10-JAN-2001; 2001US-260831P.
 PR 28-FEB-2001; 2001US-272338P.
 PR 09-MAR-2001; 2001US-274876P.
 PR 18-APR-2001; 2001US-284704P.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spylek KA, Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CM, Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grose WM, Alsodook UB, Getlach V, Edinger S, Rothenberg ME, Ellerman K, MacDougall J, Malyankar U, Miller I, Peyman J, Smithson G, Gunther E, Stone DJ;
 PI
 XX WPI: 2002-590674/63.
 DR N-PSDB; ABRN5385.
 XX
 PT NOVX polypeptides and encoding polynucleotides, useful for preventing or treating NOVX-associated disorders e.g. cancer, inflammation, or Alzheimer's disease, and in chromosome mapping, tissue typing or pharmacogenomics
 PT
 XX
 PS Claim 1; Page 57; 358bp; English.
 XX
 CC The present sequence is the protein sequence for a NOV protein. The NOV proteins and coding sequences are useful for treating or preventing CC NOV-associated disorders or in the manufacture of a medicament for treating the disorders, such as cancer, heart disease, inflammation, autoimmune disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma, IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infection (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy, and other wasting disorders associated with chronic diseases. NOV is a mixed lineage kinase 2-like protein, and the NOV coding sequence is localised to chromosome 14q24.3.
 CC
 CC
 SQ Sequence 1024 AA;
 Query Match 39.2%; Score 509.5; DB 23; Length 1024;
 Best Local Similarity 43.1%; Pred. No. 3.6e-44;
 Matches 109; Conservative 42; Mismatches 77; Indels 25; Gaps 4;
 QY 1 YRAKMTSODKEVAVK-----KLIKIKAEKILSVLSHRNIIQFYGVILEPPNY 48
 DB 147 YRAFWIGD--EVAVKARHDPDEDISQTIENVRQEKFLPMLGHPNIIALRGVCLKEPUL 204
 QY 49 GIYTEVYASUGSLVDYINSNRSEMDHMTWATVDYAKGMYLAMEAPVKVIHRLDKSN 108
 DB 205 CLVMEFARGGFLNRVLSGRK---IPPDILVNWAVQIARGNVYLHDAIVPIIHRDKSSN 261
 QY 109 VVT-----AADVGLKICDGFASRPHNHTTMSLVGTPPMKAPVVISLPSVETCDTY 160
 DB 262 VLLIQVENVGDLNSKTLKLTIDFGLAREMHTTMSAAGIYAWAPVIRASMPKSGSDVW 321
 QY 161 SYGVVLMEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEADAKR 220
 DB 322 SYGVVLMELLTGVPRFGIDGLAVAGVANKKALIPISCPSPFAKLMEDCNPPPHSR 381
 QY 221 PSFKQIISIESM 233
 DB 382 PSFTNILDQLTTI 394

RESULT 12
 ID AAE11775 standard; Protein; 1046 AA.
 XX
 AC AAE11775;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human kinase (PKIN)-9 protein.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout; cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome; AIDS; Addison's disease; microbial infection; inflammation; osteoporosis; atherosclerosis; cardiovascular disease; myocardial infarction; anaemia; myaethenia gravis; cirrhosis; cataract; growth and development disorder; seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder; lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease; obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular; antimicrobial; cytostatic; antiinflammatory; asthma.
 KW
 XX Homo sapiens.
 OS
 FH Key
 FT Location/Qualifiers
 FT 55..114
 FT /note= "SH3 domain"
 FT Domain
 FT 134..393
 FT /note= "Eukaryotic protein kinase domain"
 FT 136..386
 FT /note= "Protein kinase domain"
 FT 154..207
 FT /note= "Receptor tyrosine kinase"
 FT 181..228
 FT /note= "Receptor tyrosine kinase"
 FT 210..223
 FT /note= "Tyrosine kinase catalytic site"
 FT 232..254
 FT /note= "Receptor tyrosine kinase"
 FT 248..266
 FT /note= "Tyrosine kinase catalytic site"
 FT 290..337
 FT /note= "Receptor tyrosine kinase"
 FT 291..340
 FT /note= "Receptor tyrosine kinase"
 FT 298..330
 FT /note= "Receptor tyrosine kinase"
 FT 301..311
 FT /note= "Tyrosine kinase catalytic site"
 FT 320..342
 FT /note= "Tyrosine kinase catalytic site"
 FT 337..389
 FT /note= "Receptor tyrosine kinase"
 FT 345..389
 FT /note= "Receptor tyrosine kinase"
 FT 356..404
 FT /note= "Receptor tyrosine kinase"
 FT 364..386
 FT /note= "Tyrosine kinase catalytic site"
 XX
 PN WO200181555-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12992.
 XX
 PR 20-APR-2000; 2000US-199021P.
 PR 28-APR-2000; 2000US-200226P.
 PR 05-MAY-2000; 2000US-202339P.
 PR 11-MAY-2000; 2000US-203505P.
 PR 18-MAY-2000; 2000US-205564P.
 PR 26-MAY-2000; 2000US-207739P.
 PR 01-JUN-2000; 2000US-208795P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffen JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Redipon SA, Lu Y;
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walek R, Ramkumar J, Borowsky MD, Au-Young J, Hillman JL;
 PI Gururajan R;
 XX
 DR MPI: 2001-611740/70.
 DR N-PSDB; AAD18824.

PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders -
 XX

PS Claim 1, Page 134-136; 166pp; English.

XX The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, catarract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be
 CC in need of restorative therapy. The present sequence is human PKIN-9
 CC protein.
 CC
 XX Sequence 1046 AA;

SO Query Match 38.3%; Score 497.5; DB 22; Length 1046;

Best Local Similarity 42.3%; Pred. No. 6.9e-43;

Matches 107; Conservative 43; Mismatches 78; Indels 25; Gaps 4;

QY 1 YRAKMTSQKEVAVK-----KLIKTEKEKELISVSHRNIQFYGYLEPPY 48
 DB 149 YRAFWIGD--EVAVKARHDPDEDISQTIENVOEAKLPAMLGHPNIIALRGVCLKEPNL 206
 QY 49 GAYTEVASIGSLDYDINSNRSEMDMHTMTATDVAKGMYHMEAPYKVIHRDLKSN 108
 DB 207 CLVMEFARGGPLNRVLSGR--IPPDILVNMVQIARGNVYLDLDAIVPIIHRDLKSN 263
 QY 109 VVI-----AADGVLIKIDFGASRFHNHTMSLVGTFFMNAPEVISO LPAVSETCPTY 160
 DB 264 IILQKVENGDLSNKLKITDIFGLAREHRTTKMSAAGTANMAPEVITASMSKSDPW 323
 QY 161 SYGVAVIEMKTRFVPPKGLBGLOVAVLVEKNERLTIPSSCPSPSPALLHQCEADAKR 220
 DB 324 SYGVLLMELTGVPPRGIDGLAVAGVAMNKLALPISTCPSPFAKLMEDCWNPDPSR 383
 QY 221 PSFKOISILESM 233
 DB 384 PSFTNILDQITTI 396

RESULT 13
 ID AAE21717
 AC AAE21717 standard; Protein; 1097 AA.
 XX AAE21717;
 XX
 DT 16-JUL-2002 (first entry)
 XX

DE Human PKIN-12 protein.

XX Human, kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia;
 KW acquired immune deficiency syndrome; chronic hypoplasia; Crohn's disease;
 KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KW Down's syndrome; gene therapy; protein therapy; cytostatic.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 1..17
 FT /label= "Signal_peptide"
 FT 18..1097
 FT /note= "Mature human PKIN-12 protein"
 FT 55..114
 FT /note= "SH3 domain"
 FT 144..403
 FT /note= "Eukaryotic protein kinase domain"
 FT 146..396
 FT /note= "Protein kinase domain"
 FT 163..396
 FT /note= "Protein kinase domain"
 FT 220..233
 FT /note= "Tyrosine kinase domain"
 FT 258..276
 FT /note= "Tyrosine kinase catalytic domain"
 FT 311..321
 FT /note= "Tyrosine kinase catalytic domain"
 FT 330..352
 FT /note= "Tyrosine kinase catalytic domain"
 FT 374..396
 FT /note= "Tyrosine kinase catalytic domain"
 FT 438..749
 FT /note= "Leucine zipper domain"
 FT 869..893
 FT /note= "Leucine zipper domain"
 FT
 PN WO200218557-A2.
 XX
 PD 07-MAR-2002.
 XX

XX 31-AUG-2001; 2001WO-US27219.
 XX
 PR 31-AUG-2000; 2000US-229873P.
 PR 08-SEP-2000; 2000US-23137P.
 PR 14-SEP-2000; 2000US-232654P.
 PR 22-SEP-2000; 2000US-234902P.
 PR 29-SEP-2000; 2000US-236499P.
 PR 06-OCT-2000; 2000US-238389P.
 PR 13-OCT-2000; 2000US-240542P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.

XX Bandman O, Nguyen DB, Walla NK, Hafalia A, Yao MG, Gandhi AR;
 PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
 PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
 PI Azimati Y, Burrill JD, Marcus GA, Zingler KA, Lu DM, Lal P;
 PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
 PI Burford N;
 XX
 DR MPI: 2002-329769/36.
 DR N-PSDB; AAD34309.

XX New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as
 PT leukemia or lymphoma) -
 XX
 PS Claim 67; Page 171-173; 218pp; English.

XX The present invention relates to human kinases (PKIN) and polynucleotides

CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used
 CC in gene therapy and protein therapy. The present sequence is human
 CC PKIN-12 protein.

CC Sequence 1097 AA;

Query Match 38.3%; Score 497.5; DB 23; Length 1097;

Best Local Similarity 42.3%; Pred. No. 7.4e-43; Mismatches 25; Gaps 4;

Matches 107; Conservative 43; Indels 25; Gaps 4;

QY 1 YRAKWIISQDEKAVK-----KLLIKKEAEILSVLSHRNIOFGVILEPPNY 48
 Db 159 YRAFWIGD--EVAVKKAARHPDEDISOTIENVAQEAFLFAMLGHPNIALRGVCLKEPNL 216
 QY 49 GIVTEVASLGLDYDINSNSEEMDMHTMTATVAKGMHYLMAPVKVIRHDLKSN 108
 Db 217 CLVMEFARGGPLNRVLSGKR---IPDILVMAVQIARGMNYLLDBAIVPIIRHDLKSSN 273
 QY 109 VVI-----AADGVKICDPGASRPHNTHTMSLVGTFPMMAPIYQSLPSECTDY 160
 Db 274 ILLQKVENGLDNLKILKITDFFGLAREHRTTKMSAAGYAMMAPEVIRSMFSKSDVW 333
 QY 161 SYGVLMEMLTREVPFKLEGLOVAMLVKERNLTIPTSSCPSPFAELHQCWEADAKR 220
 Db 334 SYGVLMEMLTREVPFKLEGLOVAMLVKERNLTIPTSSCPSPFAELHQCWEADAKR 393
 QY 221 PSFKQIISLES 233
 Db 394 PSFTNILDQLTTI 406

RESULT 14

AA82886
 ID AAR82886 standard; Protein; 859 AA.

AC AAR82886;

DT 10-FEB-1996 (first entry)

DE Human leucine zipper protein-kinase.

KM Leucine zipper protein-kinase; enzyme; zpk; CNS; diagnosis;

KM central nervous system.

OS Homo sapiens.

PN W09523849-A1.

PD 08-SEP-1995.

PF 28-FEB-1995; 95WO-US02792.

PR 01-MAR-1994; 94US-0205018.

PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.

PI Pleaseure DE, Reddy U;

DR WPI; 1995-320565/41.

DR N-PSDB; AAT01031.

PT DNA encoding novel leucine zipper protein kinase - also probe(s) and
 PT polypeptide(s) prepared using the DNA, useful for inhibiting
 PT hyper-proliferation of CNS cells

PS Claim 4; Page 23-27; 40pp; English.

XX This novel protein may be used to treat tumors of the CNS in a
 CC mammal by inhibiting the overexpression of the leucine zipper
 CC protein-kinase in vivo, or by interfering with a vital signal
 CC in a chain of signals leading to tumorigenicity.

SQ Sequence 859 AA;

Query Match 37.7%; Score 490; DB 16; Length 859;

Best Local Similarity 44.3%; Pred. No. 3.2e-42; Mismatches 79; Indels 10; Gaps 6;

Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLIKIEKEI--LSVLSHRNIOFGVILEPPNYGIVTEVASLGLDYINSN 67
 Db 147 EEVAVKAVADL--KEVDIKRLKIKHPNITTFGVCQTQAPCYILMEFCQGLYELVRAG 205
 QY 68 RSEEMDMHTMTATVAKGMHYLMAPVKVIRHDLKSNVIAAGVLCIDFGASR- 126
 Db 206 RPYTPSL--LVDMSGIAGMNYLHLH--KIIRHDLKSPNMLTYDDVYKISDGTSE 260
 QY 127 FHNHTHTMSLVGTFPMMAPEVYQSLPSECTCTGSGVYLMEMLTREVPFKLEGLOVAM 186
 Db 261 LSDKSTKMSFACTVAMMAPEVIRNBPSEKVDIWSFGVLMELTGLGEIPYKVDSSAIIW 320
 QY 187 LVKERNLTIPTSSCPSPFAELHQCWEADAKKPSFKQIISLSMSNDT-SLPDK 242
 Db 321 GVGSNSLHPVSSCPDGFKILLRCQMSKPNRRSPFOILLHDLIASADVLTQOE 377

RESULT 15

AAW31227
 ID AAW31227 standard; Protein; 859 AA.

AC AAW31227;

DT 25-MAR-2003 (updated)

DT 17-FEB-1998 (first entry)

DE Human leucine-zipper protein kinase.

KM Leucine zipper protein kinase; zpk; human; metastasis; tumour;

KM serine/chreonine protein kinase; non-receptor type kinase;

OS Homo sapiens.

Key Location/Qualifiers

Domain 231..243 /label= zpk

Misc-difference 234..235 /note= "leucine zipper protein kinase"

Misc-difference 236..237 /note= "As stated in specification"

Misc-difference 240 /note= "Mentioned in specification"

Misc-difference 251 /note= "Mentioned in specification"

Misc-difference 254..256 /note= "Mentioned in specification"

Misc-difference 278..280 /note= "Mentioned in specification"

Misc-difference 292 /note= "Mentioned in specification"

Misc-difference 294..295 /note= "Mentioned in specification"

Misc-difference 297 /note= "Mentioned in specification"

Misc-difference 415..418 /note= "Mentioned in specification"

Region /note= "Putative endoplasmic reticulum targeting"

FT	Region	442..468	sequence as given in the specification"
FT		/label= "Leucine_zipper_motif	
FT		/note= "As stated in specification"	
FT	Misc-difference	443	
FT		/note= "Mentioned in specification"	
FT	Misc-difference	450	
FT		/note= "Mentioned in specification"	
FT	Misc-difference	457	
FT		/note= "Mentioned in specification"	
FT	Misc-difference	464	
FT		/note= "Mentioned in specification"	
FT	Region	537..544	
FT		/label= "ATP_binding_site	
FT		/note= "As stated in specification"	
XX	US5676945-A.		
PD	14-OCT-1997.		
XX			
PF	28-FEB-1995;	95US-0395580.	
XX			
PR	28-FEB-1995;	95US-0395580.	
XX	01-MAR-1994;	94US-0205018.	
PA	(CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.		
XX			
PI	Pleasure D, Reddy U;		
XX			
DR	WPI; 1997-511822/47.		
XX	N-PSDB; AAT89349.		
PT	Human leucine-zipper protein kinase - useful for treating tumours of		
PS	the central nervous system		
XX	Claim 3; Fig 1; 19pp; English.		
CC	This sequence represents a novel human leucine-zipper protein kinase		
CC	isolated from brain tissue. The specification states that the		
CC	protein contains an ATP-binding site at position 537-544 (consensus		
CC	sequence Gly-Xaa-Gly-Xaa-Gly), a protein kinase domain at position		
CC	231-243 and a putative endoplasmic reticulum (ER) targeting sequence at		
CC	position 413-418 (consensus sequence REEL). This protein is most similar		
CC	to members of serine/threonine protein kinases and is believed to be a		
CC	"non-receptor type kinase" based on its lack of a transmembrane domain.		
CC	Probes to this protein could be used for diagnostic or research purposes		
CC	to detect or quantitate the expression of leucine-zipper protein kinase.		
CC	Overexpression of leucine zipper protein kinase can result in		
CC	hyperproliferation of cells and metastasis. The application of exogenous		
CC	leucine-zipper protein kinase may interfere with specific protein-protein		
CC	or protein-nucleic acid interactions involved in hyperproliferation. This		
CC	may be used to treat animals suffering from tumours of the central		
CC	nervous system by inhibiting the overexpression of leucine-zipper protein		
CC	kinase in vivo or by interfering with a vital signal in the chain of		
CC	signals leading to tumourigenicity.		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
XX	Sequence	859 AA;	

Query Match	37.7%	Score 490;	DB 18;	Length 85;
Best Local Similarly	44.3%	Pred. No. 3.2e-42;		
Matches 105; Conservative	43;	Mismatches 79;	Indels 10;	Gaps 6;

QY 10 KEVAVKILKLEKEKEI--LSVLSHRNIIIOFGVILPEPNNGIYEVASLSGLYXIYNSN 67
 Db 147 EGVAVKKVRDL--KEVDIKHLKRLKRPNIITTFRGVGTQAPCYITLMEFCAQGGYIVLVRAIG 205
 QY 68 RSEEMDMDHINTWATDVAKGHHYILMEAPVKVIRHDLKSRNVIIADGVILKICDEGASR- 126
 Db 206 RPTVPSL--LVDWSMGISGAGMNYHLH--KIIRHDLKSPMNLITYDDVVKISIDPSTSKR 260
 QY 127 FHNHTHMSLVGTFPFWMAPEVYQSLPYSELCITTYSGVYLWMLREVPFKLBEGLQVAM 186

Db 261 LSDSKTNSFACTAAMMAPEVIRNEPSEKVDINSFGVLMELLTGTIPKVDVSSAIIM 320

QY 187 LVEKKERLTTPSSCPRSFAELLHQCEADAKRRPSFQIIISLESMNDT-SLPDK 242

Db 321 GVGNSNLHLPVPSSCPDEGFKILLRQCUMSKRRNRPSPFQIIHLHDIASAVLSTPQE 377

```
Search completed: December 5, 2003, 09:18:24
Job time : 27.7407 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:20:22 ; Search time 19.3519 Seconds
(without alignments)
2373.828 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277

Perfect score: 1300
Sequence: 1 YRAKWIISQKEVAVKKLTKI.....STLESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1300	100.0	455	9	US-09-757-982-5
2	1300	100.0	455	12	US-10-094-749-2477
3	1253	96.4	349	15	US-10-106-698-6345
4	510.5	39.3	394	10	US-09-862-027-19
5	497.5	38.1	1097	12	US-10-288-798-12
6	495.5	38.1	328	10	US-09-862-027-18
7	490	37.7	746	15	US-10-153-668-436
8	490	37.7	859	15	US-10-153-668-324
9	490	37.7	892	15	US-10-153-668-438
10	483	37.2	1036	12	US-10-354-358-24
11	483	37.2	1036	14	US-10-014-882-2
12	482.5	37.1	847	14	US-10-143-133-2
13	469.5	36.1	966	10	US-09-771-161A-197
14	448.5	34.5	850	10	US-09-904-389-2
15	442	34.0	263	10	US-09-840-704-5

16	442	34.0	821	12	US-10-171-404A-48	Sequence 48, Appl
17	424	32.6	92	10	US-09-764-868-799	Sequence 799, App
18	395	30.4	518	15	US-10-283-023-2	Sequence 2, Appl
19	395	30.4	579	12	US-10-384-743-4	Sequence 4, Appl
20	395	30.4	579	14	US-10-158-895-4	Sequence 15, Appl
21	395	30.4	590	12	US-10-384-743-15	Sequence 15, Appl
22	395	30.4	590	14	US-10-158-895-15	Sequence 4, Appl
23	371.5	28.6	1130	12	US-10-204-041-4	Sequence 1, Appl
24	369.5	28.4	1130	12	US-10-171-889-1	Sequence 2, Appl
25	369.5	28.4	1130	12	US-10-263-480-2	Sequence 2, Appl
26	368.5	28.3	1567	12	US-10-312-918-2	Sequence 4, Appl
27	368.5	28.3	1594	12	US-10-312-918-4	Sequence 42, Appl
28	356.5	27.4	257	12	US-09-863-776-42	Sequence 46, Appl
29	353.5	27.2	257	11	US-09-823-1874-6	Sequence 32, Appl
30	351.5	27.0	251	8	US-08-987-689A-32	Sequence 32, Appl
31	351.5	27.0	251	15	US-10-292-524-32	Sequence 41, Appl
32	345.5	26.6	252	12	US-09-976-782-41	Sequence 30, Appl
33	345.5	26.6	254	12	US-09-976-782-30	Sequence 85, Appl
34	345.5	26.6	256	12	US-09-863-776-41	Sequence 6, Appl
35	345	26.5	537	12	US-09-976-782-85	Sequence 20, Appl
36	343	26.4	505	9	US-09-977-269-6	Sequence 6, Appl
37	343	26.4	505	10	US-09-982-610-20	Sequence 40, Appl
38	343	26.4	505	10	US-09-977-260-6	Sequence 255, App
39	343	26.4	505	11	US-09-977-261-6	Sequence 2, Appl
40	341	26.2	256	12	US-09-976-782-40	Sequence 2, Appl
41	341	26.2	257	12	US-09-976-782-29	Sequence 2, Appl
42	336	25.8	1036	10	US-09-771-161A-255	Sequence 255, App
43	336	25.8	1036	10	US-09-771-161A-256	Sequence 2, Appl
44	336	25.8	1036	12	US-10-318-410-2	Sequence 5, Appl
45	336	25.8	1040	12	US-10-318-410-5	

ALIGNMENTS

RESULT 1
US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US2002094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAE-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 100.0%; Score 1300; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.8e-112;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	YRAKWIISQKEVAVKKLTKIEKEAEILSVLSRNIITQFYGVILEPPNYGLVEYASIGSL 60
DB	31	YRAKWIISQKEVAVKKLTKIEKEAEILSVLSRNIITQFYGVILEPPNYGLVEYASIGSL 90
QY	61	YDYINSRSEEMDMHIMTWADVAKGMVYLMHEAPVKVIHBDLKSNNVIAADGVLTIC 120
DB	91	YDYINSRSEEMDMHIMTWADVAKGMVYLMHEAPVKVIHBDLKSNNVIAADGVLTIC 150
QY	121	DEPASFHHHTTHMSLVGFPPMMAPEVVISLVSFSETCDTYSYGVYVMEMLTREVPKGLE 180
DB	151	DEPASFHHHTTHMSLVGFPPMMAPEVVISLVSFSETCDTYSYGVYVMEMLTREVPKGLE 210
QY	181	GLQVAVLVYKKEKRLTIPSSCFRFAELHQCWEADAKRPSFKQIISTLESMSNDTSLP 240

Db 211 GLQVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPFKQIISLESMSNDTSLP 270
QY 241 DKCNSFL 247
Db 271 DKCNSFL 277

RESULT 2

US-10-094-749-2477
; Sequence 2477, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2477
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2477

Query Match 100.0%; Score 1300; DB 12; Length 455;
Best Local Similarity 100.0%; Pred. No. 2,8e-112;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWIISQDKEVAVVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIIVTEYASLSL 60
Db 31 YRAKWIISQDKEVAVVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIIVTEYASLSL 90
QY 61 YDIINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
Db 91 YDIINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
QY 121 DFGASRFHNHTTHMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180
Db 151 DFGASRFHNHTTHMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210
QY 181 GLOVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPSPFKQIISLESMSNDTSLP 240
Db 211 GLOVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPSPFKQIISLESMSNDTSLP 270
QY 241 DKCNSFL 247
Db 271 DKCNSFL 277

RESULT 3
US-10-106-698-6345

; Sequence 6345, Application US/10106698
; Publication No. US200301096980A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6345
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (340)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6345

Query Match 96.4%; Score 1253; DB 15; Length 349;
Best Local Similarity 100.0%; Pred. No. 4.6e-108;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWIISQDKEVAVVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIIVTEYASLSL 60
Db 87 YRAKWIISQDKEVAVVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIIVTEYASLSL 146
QY 61 YDIINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
Db 147 YDIINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 206
QY 121 DFGASRFHNHTTHMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180
Db 207 DFGASRFHNHTTHMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 266
QY 181 GLOVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPSPFKQIISLESMSNDTSL 239
Db 267 GLOVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPSPFKQIISLESMSNDTSL 325

RESULT 4
US-09-862-027-19

; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

Query Match 39.3%; Score 510.5; DB 10; Length 394;
Best Local Similarity 42.7%; Pred. No. 4e-39;
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAKWIISQDKEVAVVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNY 48
||| ||| ||| : : : ||| : : |||

```

Db      18 YRAFWIGD--EVAVKAAARHDPDEDISQTIENVRQOEAKLPAMLKHPNIIALRGVCLKEPNI 75
Qy      49 GIVTEYASIGSLYDYINSNRSEEMDMHTMTATVAKGMHYLHMEAPVKVIHRDLKSRN 108
Db      76 CLVMEFARGGFLNRVLSGKR---IPPDILVNMAVQIARGMNVLHDAIVPIIHRDLKSSN 132
Qy      109 VVI-----AADGVLKICDFGASRFHNHTTMSLVGTFPMMAPEVIOSLPVSECTDPT 160
Db      133 IILQKVENGDLSNKLKLTDPGLAREMRTTKMSAAGTYAMMAPEVIRASMSFGSDVW 192
Qy      161 SYGVVIMEMLTREVPFKGLEGVQVAMLVYKNERLTIPTSCPSRFAELLHQCEADAKR 220
Db      193 SYGVLLMELTGEVPPRGIDGLRVAYGAMNKLALPIPTCPEPFKALMEDCNPDPSHR 252
Qy      221 PSFKQIISILESM 233
Db      253 PSFTNILDQLTTI 265

```

RESULT 5
US-10-288-798-12

```

; Sequence 12, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandrasekhar; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BORRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dying Aina M.; LAI, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANAVALLU, Kavitha; BURFORD, Neil;
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 55052990C01
US-10-288-798-12

```

Query Match 38.3%; Score 497.5; DB 12; Length 1097;
Best Local Similarity 42.3%; Pred. No. 2.3e-37;
Matches 107; Conservative 43; Mismatches 78; Indels 25; Gaps 4;

```

Qy      1 YRAKMSODKEVAVK-----KLLKEKEAELISVLSHRNIIQFYGVILEPPNY 48
Db      159 YRAFWIGD--EVAVKAAARHDPDEDISQTIENVRQOEAKLPAMLKHPNIIALRGVCLKEPNI 216
Qy      49 GIVTEYASIGSLYDYINSNRSEEMDMHTMTATVAKGMHYLHMEAPVKVIHRDLKSRN 108
Db      217 CLVMEFARGGFLNRVLSGKR---IPPDILVNMAVQIARGMNVLHDAIVPIIHRDLKSSN 273
Qy      109 VVI-----AADGVLKICDFGASRFHNHTTMSLVGTFPMMAPEVIOSLPVSECTDPT 160
Db      274 IILQKVENGDLSNKLKLTDPGLAREMRTTKMSAAGTYAMMAPEVIRASMSFGSDVW 333
Qy      161 SYGVVIMEMLTREVPFKGLEGVQVAMLVYKNERLTIPTSCPSRFAELLHQCEADAKR 220
Db      334 SYGVLLMELTGEVPPRGIDGLRVAYGAMNKLALPIPTCPEPFKALMEDCNPDPSHR 393
Qy      221 PSFKQIISILESM 233
Db      394 PSFTNILDQLTTI 406

```

RESULT 6
US-09-862-027-18

```

; Sequence 18, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 328
; TYPE: PRT
; ORGANISM: C. elegans
US-09-862-027-18

```

Query Match 38.1%; Score 495.5; DB 10; Length 328;

Best Local Similarity 41.0%; Pred. No. 7.9e-38;

Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

```

Qy      1 YRAKMSOD--KEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPN-YGIVTEYAS 56
Db      67 PEGNWTLPDGSQRTIALKVFVLEKEAELISKIRKNIIFGYGICATGNDFFIVTEYAF 126
Qy      57 LGSLYDYINSNRSEEM-----DMDHMTMTATVAKGMHYLHMEAPVKVIHRDLKSRN 109
Db      127 KGSLYDFHSESSQSFASSSGNSFDVYKMASQIASGIQYLHYDAVDITIIHRDLKSRN 186
Qy      110 VIAADGVKICDPGASRFHNHTTMSLVGTFPMMAPE-VIOSLPVSECTDPTSYGVVIM 167
Db      187 VLDKNLVCKICDPGSKDITHSCTAPSWGTYAMMSPEMILSGEGITLTATDWSGVVIM 246
Qy      168 EMILTREVPFKGLEGVQVAMLVYKNERLTIPTSCPSRFAELLHQCEADAKRPSFKQI 227
Db      247 EILSKVEPYKQVSEPRIFMTIQSGITLALIPSCAPLQKLSNCKMKMPKORANRQIQ 306
Qy      228 SILESMND 236
Db      307 GELNRLAGN 315

```

US-10-153-668-436

RESULT 7
US-10-153-668-436
; Sequence 436, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi

```
APPLICANT: MATSUDA, AKIO
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAMA, Kenya
TITLE OF INVENTION: STATE Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/293,172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316,031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: JP 2001-313175
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 488
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 436
LENGTH: 746
TYPE: PR1
ORGANISM: Homo sapiens
US-10-153-668-436
```

```
Query Match 37.7%; Score 490; DB 15; Length 746;
Best Local Similarity 44.3%; Pred. No. 7, 1e-37;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;
```

```
QY 10 KEVAVKVLKIKKEAEI--LSVLSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDYNIN 67
DB 180 EEVAVKVRDL-KETDIKLRKLRKHPNITTFGVCQAFCYCLMEFCAGQGLYEVLRAG 238
QY 68 RSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASR- 126
DB 239 RPTPSL--LVDMMSGIAGMNYLHLH---KIIHRDLKSPNNMLITYDDVVKISDFGTSKE 293
QY 127 FHNHTHMSLVGTFFPMMAPEVIOQLPVSECTCTYSGVGLVEMLTREVPFKGLEGLQVAM 186
DB 294 LSDKSTKMSFAGTVAMMAPEVIRNEPVSSEKVDIMSGVLMELTGEIIPYKVDSSAIIW 353
QY 187 LVEKKNERLTIPSSCFRSPFAELHOCWEADAKRPSFKOIIISLSMSNDT-SLPDK 242
DB 354 GVGSNSLHLPVSSCPDGFKILRQCWNSKPRNRPFRQILLHLDIASADVLTPOE 410
```

RESULT 8

```
US-10-153-668-324
Sequence 324, Application US/10153668
Publication No. US20030092616A1
GENERAL INFORMATION:
APPLICANT: HONDA, Goichi
APPLICANT: MATSUDA, AKIO
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAMA, Kenya
TITLE OF INVENTION: STATE Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/293,172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316,031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: JP 2001-313175
PRIOR FILING DATE: 2001-10-10
```

```
NUMBER OF SEQ ID NOS: 488
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 324
LENGTH: 859
TYPE: PR1
ORGANISM: Homo sapiens
US-10-153-668-324
```

```
Query Match 37.7%; Score 490; DB 15; Length 859;
```

```
Best Local Similarity 44.3%; Pred. No. 8, 9e-37;
```

```
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;
```

```
QY 10 KEVAVKVLKIKKEAEI--LSVLSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDYNIN 67
DB 147 EEVAVKVRDL-KETDIKLRKLRKHPNITTFGVCQAFCYCLMEFCAGQGLYEVLRAG 205
QY 68 RSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASR- 126
DB 206 RPTPSL--LVDMMSGIAGMNYLHLH---KIIHRDLKSPNNMLITYDDVVKISDFGTSKE 260
QY 127 FHNHTHMSLVGTFFPMMAPEVIOQLPVSECTCTYSGVGLVEMLTREVPFKGLEGLQVAM 186
DB 261 LSDKSTKMSFAGTVAMMAPEVIRNEPVSSEKVDIMSGVLMELTGEIIPYKVDSSAIIW 320
QY 187 LVEKKNERLTIPSSCFRSPFAELHOCWEADAKRPSFKOIIISLSMSNDT-SLPDK 242
DB 321 GVGSNSLHLPVSSCPDGFKILRQCWNSKPRNRPFRQILLHLDIASADVLTPOE 377
```

RESULT 9

```
US-10-153-668-438
Sequence 438, Application US/10153668
Publication No. US20030092616A1
GENERAL INFORMATION:
APPLICANT: HONDA, Goichi
APPLICANT: MATSUDA, AKIO
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAMA, Kenya
TITLE OF INVENTION: STATE Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/293,172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316,031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: JP 2001-313175
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 488
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 438
LENGTH: 892
TYPE: PR1
ORGANISM: Homo sapiens
US-10-153-668-438
```

```
Query Match 37.7%; Score 490; DB 15; Length 892;
```

```
Best Local Similarity 44.3%; Pred. No. 8, 9e-37;
```

```
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;
```

```
QY 10 KEVAVKVLKIKKEAEI--LSVLSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDYNIN 67
DB 180 EEVAVKVRDL-KETDIKLRKLRKHPNITTFGVCQAFCYCLMEFCAGQGLYEVLRAG 238
QY 68 RSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTICDFGASR- 126
DB 239 RPTPSL--LVDMMSGIAGMNYLHLH---KIIHRDLKSPNNMLITYDDVVKISDFGTSKE 293
```

```

QY      127 FNNHTSHSLVTFPMMAREVIOQLSPSECDYSGVGLVMEMLTREVPKRGLEGLVAM 186
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      294 LSDSTKXSFACIVAMAEVIRNEVESEAVDWMSGVVLWELLTGILPKXDWSAIIIM 353
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      187 LVEKNERLITTPSCPRSEFAELHQCMEADAKRPSFKQIISILESMNDT-SLPDX 242
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      354 GVGSNSLHLPPVSSCPDGFKILLRCQCMNSKPRRRPFROILLHDIASAVLSTRPE 410
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 10
US-10-354-358-24
; Sequence 24, Application US/10354358
; Publication No. US20030157082A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 8952, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64699, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-020PI.RNOMNIM
CURRENT APPLICATION NUMBER: US/10/354,358
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo sapiens
US-10-354-358-24

Query Match          37.2%; Score 483; DB 12; Length 1036;
Best Local Similarity 38.4%; Pred. No. 4.8e-36;
Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;

QY      1 YRAKTIQSDXEVAVKULK-----IKKAFILSVLSHRNTIQFYGVILLEPNY 48
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      139 YRATW--OQOEVAVVKARODPEDAAAAAESVRREARLFAMLRHPTIIELRGVCLOQPHL 196
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      49 GIVEVYASIGSLVDYI-----NSNRSEEDMDMIHMWTATVAGMHYLINHEAP 96
       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

Db 197 CLVFEAFAGGALNRLAALANAAPDPRAAPGRARRRIPPHVLVMAVAQIARGMYLHBEAF 256
QY 97 VKVLRDLKSNRVY-----AADGVKICDGFASRRHNHTTHMSLVSTFPMMAPEVI 148
Db 257 VPILRDLKSNILLLEKIEIHDDICNKTLLKIDFGIARERHRTTKMSTAGTYAMMAPEVI 316
QY 149 QSLPVSFETCDTYSYGVVLWEMLTREVPFKGLGLOVAMVVEKNERLTIPSSCPSPFAEL 208
Db 317 KSSLFSKSDIWSYGVLLMELLTGEVPRGIDGLAVAYGAVANVKLTPIPTSTCEPFAKL 376
QY 209 LHQCWEADAKRPPSFQIISILESMSND--TSLPDK 242
Db 377 MKECWQODPHIRPSFALLLEQLTALGAVMTMPQE 412

RESULT 11
US-10-014-882-2
; Sequence 2, Application US/10014882
; Publication No. US20020107384A1
GENERAL INFORMATION:
APPLICANT: Hu, Yi
APPLICANT: Kieke, James
APPLICANT: Donoho, Gregory
TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding th
FILE REFERENCE: LEX-0279-USA
CURRENT APPLICATION NUMBER: US/10/014,882
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: US 60/254,744
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1036
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1036)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match 37.2%; Score 483; DB 14; Length 1036;
Best Local Similarity 38.4%; Pred. No. 4,8e-36;
Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;

QY 1 YRAKVISODKEVAVAKLLK-----TEKEAEIISVLSHRNIQFYGVLLSEPPNY 48
Db 139 YRATM--QGGEVAVAAARQDPPEQDAALAAESVRREARLPFMLRHPNIIELRGVCLQQPHL 196
QY 49 GIVTEYASLGSLDYI-----NSNRSEEMDMHIMTAWDVAKGMYLMEAP 96
Db 197 CLVFEAFAGGALNRLAALANAAPDPRAAPGRARRRIPPHVLVMAVAQIARGMYLHBEAF 256
QY 97 VKVLRDLKSNRVY-----AADGVKICDGFASRRHNHTTHMSLVSTFPMMAPEVI 148
Db 257 VPILRDLKSNILLLEKIEIHDDICNKTLLKIDFGIARERHRTTKMSTAGTYAMMAPEVI 316
QY 149 QSLPVSFETCDTYSYGVVLWEMLTREVPFKGLGLOVAMVVEKNERLTIPSSCPSPFAEL 208
Db 317 KSSLFSKSDIWSYGVLLMELLTGEVPRGIDGLAVAYGAVANVKLTPIPTSTCEPFAKL 376
QY 209 LHQCWEADAKRPPSFQIISILESMSND--TSLPDK 242
Db 377 MKECWQODPHIRPSFALLLEQLTALGAVMTMPQE 412

RESULT 12
US-10-143-133-2
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
GENERAL INFORMATION:
APPLICANT: Yoganathan, Thillainathan
APPLICANT: Delaney, Allen

```

```

Db 197 CLVFEAFAGGALNRLAALANAAPDPRAAPGRARRRIPPHVLVMAVAQIARGMYLHBEAF 256
QY 97 VKVLRDLKSNRVY-----AADGVKICDGFASRRHNHTTHMSLVSTFPMMAPEVI 148
Db 257 VPILRDLKSNILLLEKIEIHDDICNKTLLKIDFGIARERHRTTKMSTAGTYAMMAPEVI 316
QY 149 QSLPVSFETCDYTSYGVVLWEMLTREVPFKGLGLOVAMVVEKNERLTIPSSCPSPFAEL 208
Db 317 KSSLFSKSDIWSYGVLLMELLTGEVPRGIDGLAVAYGAVANVKLTPIPTCPEPPAKL 376
QY 209 LHQCWEADAKRPPSFKQIISILESMSND--TSLPDK 242
Db 377 MKECWQODPHIRPSFALLLEQLTALGAVMTMPQE 412

RESULT 11
US-10-014-882-2
; Sequence 2, Application US/10014882
; Publication No. US20020107384A1
GENERAL INFORMATION:
APPLICANT: Hu, Yi
APPLICANT: Kieke, James
APPLICANT: Donoho, Gregory
TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding th
FILE REFERENCE: LEX-0279-USA
CURRENT APPLICATION NUMBER: US/10/014,882
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: US 60/254,744
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1036
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1036)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match 37.2%; Score 483; DB 14; Length 1036;
Best Local Similarity 38.4%; Pred. No. 4,8e-36;
Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;

QY 1 YRAKVISODKEVAVKLLK-----TEKEAELISVLSHRNIQFYGVLLSEPPNY 48
Db 139 YRATM--QGGEVAVAAARQDPPEQDAALAAESVRERARLPFMLRHNPITIELRGVCLQCPHL 196
QY 49 GIVTEYASLGSLDYI-----NSNRSEEMDMHIMTAWDVAKGMYLMEAP 96
Db 197 CLVFEAFAGGALNRLAALANAAPDPRAAPGRARRRIPPHVLVMAVAQIARGMYLHBEAF 256
QY 97 VKVLRDLKSNRVY-----AADGVKICDGFASRRHNHTTHMSLVSTFPMMAPEVI 148
Db 257 VPILRDLKSNILLLEKIEIHDDICNKTLLKIDFGIARERHRTTKMSTAGTYAMMAPEVI 316
QY 149 QSLPVSFETCDYTSYGVVLWEMLTREVPFKGLGLOVAMVVEKNERLTIPSSCPSPFAEL 208
Db 317 KSSLFSKSDIWSYGVLLMELLTGEVPRGIDGLAVAYGAVANVKLTPIPTCPEPPAKL 376
QY 209 LHQCWEADAKRPPSFKQIISILESMSND--TSLPDK 242
Db 377 MKECWQODPHIRPSFALLLEQLTALGAVMTMPQE 412

RESULT 12
US-10-143-133-2
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
GENERAL INFORMATION:
APPLICANT: Yoganathan, Thillainathan
APPLICANT: Delaney, Allen

```

TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
FILE REFERENCE: KIN-023
CURRENT APPLICATION NUMBER: US/10/143,133
CURRENT FILING DATE: 2002-05-09
PRIOR APPLICATION NUMBER: 60/290,555
PRIOR FILING DATE: 2001-05-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 847
TYPE: PRT
ORGANISM: Homo sapien
US-10-143-133-2

Query Match 37.1%; Score 482.5; DB 14; Length 847;
Best Local Similarity 41.1%; Pred. No. 4,1e-35;
Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;

QY 1 YRAKWSQ-----DKKVAVKLLKIEKAEILSVSHRNIIQFYGVILEPPNYG 49
DB 132 YRSGWKGELVAVKARODPEDISV-TAESVROEARLFAHLAHPNIIALKAACLEEFNLC 190
QY 50 IYVEYASLSLYDYNINSRSEMDMDHIM-TWATDVAKGMHYLHMEAPVKVIRHDLKSRN 108
DB 191 LVMEYVAGGFLSRALAGRVP---PHVLVMAVQIARGMHYLHCEALVPIVIRHDLKSN 246
QY 109 VV-----IADGV---LKICDFGASRPHNHTHMSLVGTTPMMAPEVIQSLPVSETCDTY 160
DB 247 ILLQPIESDDMEHKTKITDPGLAREMHKTQMSAQTAMMAPEVIKASTFSKSDVW 306
QY 161 SYGVLMEMLTREVPFKGLEQVAVLVEKNEELTIPSSCPSPFAELHQCWEADAKKR 220
DB 307 SFGVLMELTGEVTPFGICDLAVAVGVAVVANKLTPIPTCPPEFQALMADCAQDPHRR 366
QY 221 PSFKQIISILESM 233
DB 367 PDFASLIQLEAL 379

RESULT 13
US-09-771-161A-197
Sequence 197, Application US/09771161A
Patent No. US20020110811A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724,676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 197
LENGTH: 966
TYPE: PRT
ORGANISM: Homo sapiens
US-09-771-161A-197

Query Match 36.1%; Score 469.5; DB 10; Length 966;
Best Local Similarity 42.2%; Pred. No. 7.8e-35;
Matches 97; Conservative 47; Mismatches 77; Indels 9; Gaps 5;

QY 10 KEVAVKLLKIEKAEI--LSVSHRNIIQFYGVILEPPNYGVITEVASLSLYDYN 67
DB 190 EEVAIKK-VBEQNETDIKHRLKRGHPNIIAFKGVCTQAPCYCIIMECAHGOLEYEVLRA 248
QY 68 RSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVIIADGVAKICDFGASR- 126

DB 249 R--KITPRLLVDMSTGSIAGMNVYLH---KIIHRDLKSPNVLVHTTDAVKISDFGTSK 303
QY 127 FHNHTHMSLVGTTPMMAPEVIQSLPVSETCDTYSYGVLMEMLTREVPFKGLEQVAM 186
DB 304 LSDKSTKMSFAGTAVMAAPEVIRNEFVSKEVIMSGVGLWELLTGEIYKOVDSASIIW 363
QY 187 LVYKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMND 236
DB 364 GVGSNSLHLPVPSITCPDGFKILMKQITWQSKPRNRPSPFQTLHMLDIAAD 413

RESULT 14
US-09-904-389-2
Sequence 2, Application US/09904389
Patent No. US20020129404A1
GENERAL INFORMATION:
APPLICANT: Clendennen, Stephanie K.
APPLICANT: Schuster, Debra K.
TITLE OF INVENTION: CTRL HOMOLOGUE FROM MELON
FILE REFERENCE: 4257-0029.30
CURRENT APPLICATION NUMBER: US/09/904,389
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/218,307
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 850
TYPE: PRT
ORGANISM: Cucumis melo
FEATURE:
NAME/KEY: VARIANT
LOCATION: (154)...(154)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

Query Match 34.5%; Score 448.5; DB 10; Length 850;
Best Local Similarity 40.2%; Pred. No. 5.9e-33;
Matches 99; Conservative 38; Mismatches 94; Indels 15; Gaps 4;

QY 1 YRAKWSIQDEKVAVKLLKIE-----KEAEILSVSHRNIIQFYGVILEPPNYGI 50
DB 592 YRGEWHGSD--VAVKILTBQDFPERVNEFLREVAIMKSLRHPNIIYLFMGAVTKPPNLSI 649
QY 51 YRYASLSLYDYNINSRSEMDMDHIMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNV 110
DB 650 YVEYLSRGSLLYRLHRSKGVKIDETRIIMAFDVAKGMVYLHRRDP-PIVHDLKSPNLL 708
QY 111 IADGVAKICDFGASRPHNHT--THMSLVGTTPMMAPEVIQSLPVSETCDTYSYGVLM 168
DB 709 VDKKTVKVCDFGLSLKARTPLSSKSAAGTERMAPEVLRRDEPSNEKSDVYSFGVILME 768
QY 169 MLTREVPFKGLEQVAVLVEKNEELTIPSSCPSPFAELHQCWEADAKKRPSFKQIIS 228
DB 769 LATIQPCWNLNPAQVAVAVGFKRGLDIPDVNPKLASLIYACWADPEMKSPSSISME 828
QY 229 ILESMS 234
DB 829 TLKPMT 834

RESULT 15
US-09-840-704-5
Sequence 5, Application US/09840704
Patent No. US20020122801A1
GENERAL INFORMATION:
APPLICANT: Dedhar, Shoukat
APPLICANT: Hamigan, Greg
TITLE OF INVENTION: Interglin-Linked Kinase and Its Uses
FILE REFERENCE: KIN-2CON
CURRENT APPLICATION NUMBER: US/09/840,704
CURRENT FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/566,906

; PRIOR FILING DATE: 2000-05-09
 ; PRIOR APPLICATION NUMBER: US08/752,345
 ; PRIOR FILING DATE: 1996-11-19
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 263
 ; TYPE: PRF
 ; ORGANISM: H. sapiens
 ; FEATURE:
 ; NAME/KEY: Other
 ; LOCATION: (1)...(263)
 US-09-840-704-5

Query Match 34.0%; Score 442; DB 10; Length 263;
 Best Local Similarity 40.2%; Pred. No. 5.5e-33;
 Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;

QY	1	YRAKNTSODKEVAVKKLKLIE-----KEAETLSVLSHRNITQFYGVILEPPNYGI	50
DB	21	HRAEMHGSD--VAVKILMEODFHAERVNEFLREVAIMKRLRHPNIVLEMGAVTQPPNLSI	78
QY	51	VTEYASLGSLYDYIN--SNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV	109
DB	79	VTEYLSRGLYRLHLSGARBEQDERRRISMAYDAKGMNYLHNRNP-PIVHRDLKSPNL	137
QY	110	VIAADGVLCIDPGASRFNNHT--THMSLVGTFPMMAPEVYOSLPVSETCDTYSYGVILM	167
DB	138	LVDKKTYVAVCDPGLSLKASTFLSSKSAAGTPPEMMAPEVLRDEPSNEKSDVYSFGVILM	197
QY	168	EMLTREVPFKGLEGLQVAMLVVEKNERLITPSCCPSPAFELHQWEADAKKRPSPFOJII	227
DB	198	ELATLQOPWGNLNPQAVVAAGVCKRLEIPRNINPQVAALIGCWNTNEPWRKRPSPATTM	257
QY	228	SILESMT233	
DB	258	DLRPL263	

Search completed: December 5, 2003, 09:29:17
 Job time : 20.3519 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:17:01 ; Search time 10.5556 Seconds
(without alignments)
990.074 Million cell updates/sec

Title: US-09-757-982-5_COPY_31_277
Perfect score: 1300
Sequence: 1 YRAKWIISQDKEVAVKKLKI.....STLEGSNDTSLPDKNSFL 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1300	100.0	455 3	US-09-221-235-5 Sequence 5, Appl1
2	1300	100.0	455 3	US-09-221-928-5 Sequence 5, Appl1
3	1300	100.0	455 3	US-09-221-527-5 Sequence 5, Appl1
4	1300	100.0	455 3	US-09-221-236-5 Sequence 5, Appl1
5	1300	100.0	455 3	US-09-221-416-5 Sequence 5, Appl1
6	1300	100.0	455 3	US-09-221-245-5 Sequence 5, Appl1
7	1300	100.0	455 3	US-09-163-115-5 Sequence 5, Appl1
8	1300	100.0	455 3	US-09-221-528-5 Sequence 5, Appl1
9	1300	100.0	455 3	US-09-593-553-5 Sequence 5, Appl1
10	1300	100.0	455 3	US-09-221-237-5 Sequence 5, Appl1
11	1300	100.0	455 4	US-09-399-588-2 Sequence 2, Appl1
12	510.5	39.1	394 4	US-09-345-473B-19 Sequence 18, Appl1
13	495.5	38.1	328 4	US-08-395-580-2 Sequence 2, Appl1
14	490	37.7	668 1	US-08-395-580-2 Sequence 2, Appl1
15	490	37.7	859 1	PCT-US95-02792-2 Sequence 2, Appl1
16	490	37.7	859 5	PCT-US95-02792-2 Sequence 2, Appl1
17	442	34.0	263 3	US-09-035-706-5 Sequence 5, Appl1
18	442	34.0	263 3	US-08-955-841-5 Sequence 5, Appl1
19	442	34.0	263 4	US-09-390-425-5 Sequence 5, Appl1
20	442	34.0	263 4	US-09-566-906-5 Sequence 5, Appl1
21	442	34.0	821 1	US-07-928-464-2 Sequence 2, Appl1
22	442	34.0	821 1	US-08-003-311B-2 Sequence 2, Appl1
23	442	34.0	821 1	US-08-261-432-2 Sequence 2, Appl1
24	442	34.0	821 5	PCT-US93-07347-2 Sequence 2, Appl1
25	395	30.4	579 4	US-09-529-279-4 Sequence 4, Appl1
26	395	30.4	579 4	US-10-158-895-4 Sequence 4, Appl1
27	395	30.4	590 4	US-09-529-279-5 Sequence 15, Appl1

28	395	30.4	590 4	US-10-158-895-15 Sequence 15, Appl1
29	363.5	28.0	275 2	US-08-701-191A-36 Sequence 36, Appl1
30	359.5	27.7	261 2	US-07-857-224B-59 Sequence 59, Appl1
31	356.5	27.4	261 2	US-07-857-224B-60 Sequence 60, Appl1
32	355.5	27.3	1584 4	US-09-457-040B-27 Sequence 27, Appl1
33	343	26.4	505 1	US-08-222-616-20 Sequence 20, Appl1
34	343	26.4	505 4	US-08-446-648-20 Sequence 20, Appl1
35	343	26.4	505 4	US-08-492-723-2 Sequence 2, Appl1
36	343	26.4	505 5	PCT-US95-04228-20 Sequence 20, Appl1
37	341	26.2	506 4	US-08-426-509A-6 Sequence 6, Appl1
38	341	26.2	511 4	US-08-232-545-6 Sequence 6, Appl1
39	341	26.2	511 5	PCT-US95-05008-6 Sequence 6, Appl1
40	339	26.1	820 1	US-08-166-717D-6 Sequence 6, Appl1
41	335.5	25.8	259 2	US-07-857-224B-52 Sequence 52, Appl1
42	335.5	25.8	537 4	US-08-426-509A-11 Sequence 11, Appl1
43	335.5	25.8	537 4	US-08-232-545-11 Sequence 11, Appl1
44	335.5	25.8	537 5	PCT-US95-05008-11 Sequence 11, Appl1
45	335.5	25.8	835 4	US-09-458-457-8 Sequence 8, Appl1

ALIGNMENTS

```
RESULT 1
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match      100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YRAKWIISQDKEVAVKKLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASIGSL 60
Db      |||
QY      31 YRAKWIISQDKEVAVKKLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASIGSL 90
Db      |||
QY      61 YYINSNRSEENDMDHIMTATDVAKGMAYLHMEAPVKVIHRDLKSRNVVIADGVLTIC 120
Db      |||
QY      91 YYINSNRSEENDMDHIMTATDVAKGMAYLHMEAPVKVIHRDLKSRNVVIADGVLTIC 150
Db      |||
QY      121 DPGASRFHHHTHTMSLVGTFPMWAPRVIOSLPVSEICDTYSGVVLMEMLTRVPPKGL 180
Db      |||
QY      151 DPGASRFHHHTHTMSLVGTFPMWAPRVIOSLPVSEICDTYSGVVLMEMLTRVPPKGL 210
Db      |||
QY      181 GQVAMLVVEKNERLTPSSCPSPFAELHQCWEADAKRPSFKOISILEGNSNTSLP 240
Db      |||
QY      211 GQVAMLVVEKNERLTPSSCPSPFAELHQCWEADAKRPSFKOISILEGNSNTSLP 270
Db      |||
QY      241 DKCNSTL 247
Db      |||
QY      271 DKCNSTL 277

RESULT 2
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
```

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PR1
ORGANISM: Homo sapiens
US-09-221-928-5

Query Match 100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 60
1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 90
61 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
91 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
121 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180
151 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210
181 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 240
211 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 270
241 DKNSFL 247
271 DKNSFL 277

RESULT 3
US-09-221-527-5
Sequence 5, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PR1
ORGANISM: Homo sapiens
US-09-221-527-5

Query Match 100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 60
1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 90
61 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
91 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
121 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180

151 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210
181 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 240
211 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 270
241 DKNSFL 247
271 DKNSFL 277

RESULT 4
US-09-221-236-5
Sequence 5, Application US/09221236
Patent No. 6146841
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,236
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PR1
ORGANISM: Homo sapiens
US-09-221-236-5

Query Match 100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 60
1 YRAKMSODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSL 90
61 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120
91 YDYINSRSEEMDMHMTATVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150
121 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180
151 DFGASRFHNHTTMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210
181 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 240
211 GIOVAMLVVEKNERLTPSSCFPSFAELHQCWEADAKKRPFSKQIISLESMSNDTSLP 270
241 DKNSFL 247
271 DKNSFL 277

RESULT 5
US-09-221-416-5
Sequence 5, Application US/09221416
Patent No. 6153417
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,416
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match
Best Local Similarity 100.0%; Score 1300; DB 3; Length 455;
Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 60
   |||
DB 31 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 90

QY 61 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
   |||
DB 91 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150

QY 121 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
   |||
DB 151 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 240
   |||
DB 211 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
   |||
DB 271 DKCNSFL 277

RESULT 6
US-09-221-245-5
; Sequence 5, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-5

Query Match
Best Local Similarity 100.0%; Score 1300; DB 3; Length 455;
Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 60
   |||
DB 31 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 90

QY 61 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
   |||
DB 91 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150

QY 121 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
   |||
DB 151 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 240
   |||
DB 211 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
   |||
DB 271 DKCNSFL 277
```

```

RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match
Best Local Similarity 100.0%; Score 1300; DB 3; Length 455;
Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 60
   |||
DB 31 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 90

QY 61 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
   |||
DB 91 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150

QY 121 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
   |||
DB 151 DFGASRFHNTHTMSLVGTFFPMMAPEVIOSLPVSETCDITYSGVVLWEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 240
   |||
DB 211 GLOVAMLVVEKNERLITPSSCPSPFAELHQCWEADAKKRPSPKQIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
   |||
DB 271 DKCNSFL 277

RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

Query Match
Best Local Similarity 100.0%; Score 1300; DB 3; Length 455;
Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 60
   |||
DB 31 YRAKWSODKEVAVKKLLIKEAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLGS 90

QY 61 YDYINSRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
```

```

; DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 150
;
; 121 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
;
; DB 151 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 210
;
; QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 240
;
; DB 211 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 270
;
; QY 241 DKCNSTL 247
;
; DB 271 DKCNSTL 277

```

```

RESULT 9
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPE-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

```

```

Query Match 100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 60
;
; DB 31 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 90
;
; QY 61 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 120
;
; DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 150
;
; QY 121 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
;
; DB 151 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 210
;
; QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 240
;
; DB 211 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 270
;
; QY 241 DKCNSTL 247
;
; DB 271 DKCNSTL 277

```

```

RESULT 10
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPE-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; PRIOR FILING DATE: 1998-12-28
; PRIOR FILING DATE: 09/163,115
; EARLIER FILING DATE: 1998-09-29

```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

```

```

Query Match 100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 60
;
; DB 31 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 90
;
; QY 61 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 120
;
; DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 150
;
; QY 121 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
;
; DB 151 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 210
;
; QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 240
;
; DB 211 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 270
;
; QY 241 DKCNSTL 247
;
; DB 271 DKCNSTL 277

```

```

RESULT 11
US-09-399-588-2
; Sequence 2, Application US/09399588
; Patent No. 6511825
; GENERAL INFORMATION:
; APPLICANT: Rugieri, Rosamaria
; APPLICANT: Callow, Marinella
; TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids
; FILE REFERENCE: 1044-US
; CURRENT APPLICATION NUMBER: US/09/399,588
; PRIOR FILING DATE: 1999-09-20
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Human J42
US-09-399-588-2

```

```

Query Match 100.0%; Score 1300; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 60
;
; DB 31 YRAKWSODKEVAVKKLTIKEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLSGL 90
;
; QY 61 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 120
;
; DB 91 YDINSNRSEMDMDHMTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLTIC 150
;
; QY 121 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 180
;
; DB 151 DFGASRFHNHTHTMSLVGTFPPMAPEVIOQLPVSETCDITYSGVVLWEMLTREVPFKGLE 210
;
; QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKKRPFKQIISLESMSNDTSLP 240
;

```

```

Db      211 GLQVAMLVVEKNERLTIIPSSCPSPFAELLHQCEWADAKRRPSFKQIISLESWSDTSLP 270
Qy      241 DKCNSTL 247
Db      271 DKCNSTL 277

RESULT 12
US-09-345-473E-19
; Sequence 19, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-345-473E-19

Query Match      39.3%; Score 510.5; DB 4; Length 394;
Best Local Similarity 42.7%; Pred. No. 2e-45;
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

Qy      1 YRAKWSODKEVAVK-----KLIKIEKAEILSVLSHRNIQFVGLLEPPNY 48
Db      18 YRAFWIGD--EVAVKARHDPDEDISQTIENVAGKLFAMLKHPIIALRGCLKEPNI 75
Qy      49 GIVTEYASIGSLYDYINNSNSEBMDMHTWATDVAKGMYLHMEAPVKVTHRDLSRN 108
Db      76 CLVMEFARGGPLNRVLSGKR---IPPDILVMAVQIARGMNYLHDAIVPIIHRDLKSNV 132
Qy      109 VVI-----ADGVLIKICDGFASRFHNHTTMSLVGTFFPMWAPVYIOSLPVSECDPY 160
Db      133 ILILQVENGDDLSNKLTKITDPLARBMHRTTMSAAGIYAMWAPVIRASMPKSDVW 192
Qy      161 SYGVVLMEMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELLHQCEWADAKRR 220
Db      193 SYGVLLMELLTGVRPFRGIDGLRVAAGVANMKLALPIPTCPEPFAKLMDCMNPDPHNR 252
Qy      221 PSFKQIISLESW 233
Db      253 PSFTNILDQLTTI 265

RESULT 13
US-09-345-473E-18
; Sequence 18, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 328
; TYPE: PRT
; ORGANISM: C. elegans
US-09-345-473E-18

Query Match      38.1%; Score 495.5; DB 4; Length 328;
Best Local Similarity 41.0%; Pred. No. 5.9e-44;
Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

Qy      1 YRAKWSOD---KEVAVKLLKIEKAEILSVLSHRNIQFVGLLEPPN-YGIVTEYAS 56

```

```

Db      67 FSGNWTLPDGSQRTIALKKVFVLEKEAEILSKIRHKNIIQFVGICATGNDFFIVTEYAE 126
Qy      57 LQSLDYINSNRSEEM-----DMDHMTWATDVAKGMYLHMEAPVKVTHRDLSRN 109
Db      127 KGLYDFFIHSESSQSPASSGNSPFWVKMASQIASGIQYLHYDAVDPIIHRDLKSKNV 186
Qy      110 VIAADGVLIKICDGFASRFHNHT--THMSLVGTFFPMWAPV-VIOSLPVSECDPYSGVVLW 167
Db      187 VLDKRLVCKICDGFSGTSLTHSCTAPSWGGIAMSPEWITLQSEGLTTTWDVWSYGVVLW 246
Qy      168 EMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELLHQCEWADAKRRPSFKQII 227
Db      247 ELTSKEVPYKDXSEPRIFPMITQSGITLAIPTSCAPALQMSNCKMTPKORANNRQIQ 306
Qy      228 STLESWSND 236
Db      307 GEINRLAGN 315

RESULT 14
US-08-205-018-2
; Sequence 2, Application US/08205018
; Patent No. 5554523
; GENERAL INFORMATION:
; APPLICANT: Reddy, Usharani R.
; TITLE OF INVENTION: Pleasure, David
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; ADDRESSEE: No. 5554523rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,018
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaumond, Rebecca R.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-205-018-2

Query Match      37.7%; Score 490; DB 1; Length 668;
Best Local Similarity 44.3%; Pred. No. 6.2e-43;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

Qy      10 KEVAVKLLKIEKAEI--LSVLSHRNIQFVGLLEPPNYGIVTEYASLSGLSYDINSN 67
Db      147 EEVAVKVRVDL-KETDIKRLRKIKHPNITTFGVCTQACYCCLMEFCAQGGQLYEVLRAQ 205
Qy      68 RSEBMDHMTWATDVAKGMYLHMEAPVKVTHRDLSRNVIADGVLIKICDGFASR- 126

```

Db 206 RPTPSL--LVDSMSIAGNNYLHH--KIIHDELSPPMLITDDVVKISPGFSKE 260

QY 127 FHHHTHMSLVGTFPMNAPEVIOSLPVSECTGYSYGAVLWEMLTREVPFGEGGLQVAM 186

Db 261 LSPKSTKMSFAGVAMAAPEVIRNEVSVSKVDIWSFGVVLWELLTGELPIYDVDDSSAIIV 320

QY 187 LVYENKERLTIPESSCPSPFAELLHOCWEDADKKRSPFOIISLESMSNDI-SLPDK 242

Db 321 GVGSNLSHLPEVPSKCDGFKILRLQCSMSKPEPNRSPFOILLHLDIASADVLSLPOE 377

Db 321 GVGSNLSLHPVPSSCPFGFKILLRQCWNSKPRNRPSEFRQILLHLDIASADVLSTPQE 377

```
Search completed: December 5, 2003, 09:21:41
Job time : 11.5556 secs
```

```

1      RESULT 15
2      US-08-395-580-2
3      : Sequence 2, Application us/08395580
4      : Parent No. 5676945
5      : GENERAL INFORMATION:
6      : APPLICANT: Ushaarani R. Reddy, David Pleasure and the Children's
7      : APPLICANT: Hospital of Philadelphia
8      : TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
9      : TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
10     : NUMBER OF SEQUENCES: 4
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945f1s
13     : STREET: One Liberty Place - 46th Floor
14     : CITY: Philadelphia
15     : STATE: PA
16     : COUNTRY: U.S.A.
17     : ZIP: 19103
18     : COMPUTER READABLE FORM:
19     : MEDIUM TYPE: 3.5 inch disk, 720 Kb
20     : COMPUTER: IBM PC compatible
21     : OPERATING SYSTEM: PC-DOS/MS-DOS
22     : SOFTWARE: Wordperfect 5.1
23     : CURRENT APPLICATION DATA:
24     : APPLICATION NUMBER: US/08/395,580
25     : FILING DATE: herewith
26     : CLASSIFICATION: 514
27     : PRIOR APPLICATION DATA:
28     : APPLICATION NUMBER: 08/205,018
29     : FILING DATE: 01-MAR-1994
30     : ATTORNEY/AGENT INFORMATION:
31     : NAME: Rebecca L. Ralph (formerly Gaumond)
32     : REGISTRATION NUMBER: 35,152
33     : REFERENCE/DOCKET NUMBER: CH-0488
34     : TELECOMMUNICATION INFORMATION:
35     : TELEPHONE: 215-568-3100
36     : TELEFAX: 215-568-3439
37     : INFORMATION FOR SEQ ID NO: 2:
38     : SEQUENCE CHARACTERISTICS:
39     : LENGTH: 859 amino acids
40     : TYPE: amino acid
41     : TOPOLOGY: linear
42     : MOLECULE TYPE: protein
43     : US-08-395-580-2

```

Query Match	37.7%	Score 490;	DB 1;	Length 859;
Post Local similarity	44.3%	Prod No. 9	00-43	

QY 10 KEVAVKLLKLEKEKEI--LSVLSHRNIIQFGLVLEPPNNGIVTEVYASLSIDVYINSN 67
 Db 147 EGVAVKKVRDL-KENDIDKHLRKLKHPNIIITFGVCTQAPCYCLIMEFCAQGLVEVLVLAG 205
 QY 68 RSEEMDMHITWATDVAKGMHYIHLMEAPVKYIHDLKSRNVLIADVLKICDFGAR- 126
 Db 206 RPVTPSL--LVDMSWGIAGGMNYLHLH--KIIRHDLKSPNNLLITDYDVVAKISDGTSK 260
 QY 127 FHNHTTHSLVGFPEWMAPEVYIOSLPVSETCDTYSYGVVLWEMLTREVPFKLEGLAQVAM 186
 Db 261 LSDSTKNSPAGTVAMAPAEVIRNRPVSEKNDVMSFGVVLWELLGELPYKDVOSSALII 320
 QY 187 LVEKNEMLTIPSSCPKSPFAELLHOCWEADAKKRPSPFOITISLMSNDT-SLEPK 242

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:28:12 ; Search time 21 Seconds
(without alignments)
2083.655 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSLSASFQIKFDLQFFE.....GDDDDDDGGEEDNDMNSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

1: PIR76:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	2.4	848	2 B87950	protein F33E2.2 [1
2	11	2.4	855	2 T20082	hypothetical prote
3	10	2.2	729	2 A29651	KEX1 protein precu
4	10	2.2	1089	2 S48244	NMD2 protein - yea
5	9	2.0	297	2 T30613	hypothetical prote
6	9	2.0	375	2 A39622	protein kinase MCK
7	9	2.0	408	2 T08069	protein kinase, 48
8	9	2.0	502	2 JC4337	activin receptor 1
9	9	2.0	502	2 I48241	ALK-1 - mouse
10	9	2.0	764	2 S64951	hypothetical prote
11	9	2.0	856	2 T52415	polycomb protein B
12	9	2.0	898	2 T01503	hypothetical prote
13	9	2.0	1363	2 T43320	insulin-like growt
14	9	2.0	1390	2 T30346	insulin receptor -
15	9	2.0	2391	2 T18410	carbamoyl-phosphat
16	8	1.8	101	2 T38320	protein-serine/thr
17	8	1.8	108	2 T38312	protein-serine/thr
18	8	1.8	157	2 S09814	hypothetical prote
19	8	1.8	158	1 T05710	2S albumin precurs
20	8	1.8	162	2 B81330	hypothetical prote
21	8	1.8	198	2 A96520	hypothetical prote
22	8	1.8	211	2 D69413	conserved hypotet
23	8	1.8	229	1 P08R82	antiterminal Q -
24	8	1.8	229	2 D90831	antiterminal pr
25	8	1.8	229	2 G85688	probable antiterm
26	8	1.8	237	2 T47271	hypothetical prote
27	8	1.8	250	2 T47611	hypothetical prote
28	8	1.8	260	2 G84645	hypothetical prote
29	8	1.8	286	2 A34599	DNA-binding protei

30	8	1.8	294	1 A37818	osteopontin precu
31	8	1.8	298	2 T52117	zinc finger protei
32	8	1.8	301	2 TC5811	osteopontin - rat
33	8	1.8	317	1 A25917	osteopontin precu
34	8	1.8	318	2 E69026	conserved hypotet
35	8	1.8	328	2 T16747	hypothetical prote
36	8	1.8	352	1 R5HSL0	ribosomal protein
37	8	1.8	352	2 G84266	50S ribosomal prot
38	8	1.8	368	2 G96668	protein FIN19.7 (i
39	8	1.8	374	2 T33173	hypothetical prote
40	8	1.8	379	2 T04645	hypothetical prote
41	8	1.8	390	2 T01451	protein kinase hom
42	8	1.8	391	2 D64366	hypothetical prote
43	8	1.8	422	2 S52578	serine/threonine-s
44	8	1.8	426	2 T05676	hypothetical prote
45	8	1.8	434	2 B54843	nemo, form II - fr
46	8	1.8	436	2 T51402	serine/threonine-s
47	8	1.8	439	2 T31734	hypothetical prote
48	8	1.8	471	2 T39232	probable serine th
49	8	1.8	477	2 A54843	nemo, form I - fru
50	8	1.8	482	2 S37845	transcription init
51	8	1.8	485	2 T49237	hypothetical prote
52	8	1.8	494	2 D84860	probable protein k
53	8	1.8	504	2 T10558	hypothetical prote
54	8	1.8	513	2 B96524	hypothetical prote
55	8	1.8	526	2 T47786	hypothetical prote
56	8	1.8	538	2 C83284	probable biotin-de
57	8	1.8	541	2 A12238	glutamine-fructose
58	8	1.8	560	2 T49839	hypothetical prote
59	8	1.8	567	2 JC5957	transforming growt
60	8	1.8	579	2 JC5955	transforming growt
61	8	1.8	584	2 T48206	phosphoinositide-s
62	8	1.8	590	2 S57594	hypothetical prote
63	8	1.8	604	2 JC5956	transforming growt
64	8	1.8	646	2 C96596	hypothetical prote
65	8	1.8	650	2 C96572	protein Fl2M16.4 l
66	8	1.8	746	2 G02838	enhancer-of-zeste
67	8	1.8	800	2 T02852	probable membrane
68	8	1.8	836	2 B66716	probable serine/th
69	8	1.8	840	2 S48975	hypothetical prote
70	8	1.8	847	1 A53800	mixed-linase prot
71	8	1.8	872	2 T18861	probable peptide-a
72	8	1.8	938	2 A56731	chromatin assembly
73	8	1.8	943	2 B45082	neurotrophic recep
74	8	1.8	992	2 T05335	hypothetical prote
75	8	1.8	992	2 S49835	hypothetical prote
76	8	1.8	1019	2 T13039	lysine kinase re
77	8	1.8	1038	2 T02634	rep protein homolo
78	8	1.8	1085	2 S55352	IFH protein - yea
79	8	1.8	1128	2 G86266	hypothetical prote
80	8	1.8	1131	2 S22266	FUN30 protein - ye
81	8	1.8	1165	2 S62982	vacuolar protein V
82	8	1.8	1234	2 T00363	hypothetical prote
83	8	1.8	1341	2 S66835	hypothetical prote
84	8	1.8	2251	2 T24490	probable membrane
85	8	1.8	3394	2 T18501	hypothetical prote
86	8	1.5	51	2 P0479	cell division cycl
87	8	1.5	54	2 S56724	abscisic acid-indu
88	8	1.5	54	2 S56725	conserved hypotet
89	8	1.5	71	2 A69154	hypothetical prote
90	8	1.5	85	2 F97131	hypothetical prote
91	8	1.5	95	2 T46086	hypothetical prote
92	8	1.5	101	1 R36CCK	transcription repr
93	8	1.5	107	2 T35523	probable small sec
94	8	1.5	108	2 T16893	hypothetical prote
95	8	1.5	111	2 T38218	protein-serine/thr
96	8	1.5	114	2 D46516	Ig lambda chain V
97	8	1.5	116	2 AD2693	conserved hypotet
98	8	1.5	123	2 T49794	hypothetical prote
99	8	1.5	133	2 T46978	hypothetical prote
100	8	1.5	133	2 A81435	PTS system, fructo

ALIGNMENTS

RESULT 1
 B87950
 Protein F33E2.2 [imported] - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C/Accession: B87950
 R/Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A/Reference number: A75000; MUID:99069613; PMID:9851916
 A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A/Accession: B87950
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-848 <STO>
 A/Cross-references: GB:chr_I; PIDN:CAB06544.1; PID:g3876653; GSPDB:GN00019; CESP:F33E2.2
 C/Genetics:
 A/Map position: 1
 A/Map position: 1

Query Match 2.4%; Score 11; DB 2; Length 848;
 Best Local Similarity 100.0%; Pred. No. 0.044;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GVVLMWLTRE 203
 DB 235 GVVLMWLTRE 245

RESULT 2

T20082
 hypothetical protein F33E2.2 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T20082; T21703
 R/Lennard, N.
 submitted to the EMBL Data Library, April 1998
 A/Reference number: Z19221
 A/Accession: T20082
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-855 <WIL>
 A/Cross-references: EMBL:AL022593; PIDN:CAA18635.2; GSPDB:GN00019; CESP:F33E2.2
 A/Experimental source: clone C49C9
 R/Lennard, N.
 submitted to the EMBL Data Library, January 1997
 A/Reference number: Z19461
 A/Accession: T21703
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-855 <WIL>
 A/Cross-references: EMBL:Z84574; PIDN:CAB06544.2; GSPDB:GN00019; CESP:F33E2.2
 A/Experimental source: clone F33E2
 C/Genetics:
 A/Map position: 1
 A/Map position: 1

QY 193 GVVLMWLTRE 203
 DB 235 GVVLMWLTRE 245

Query Match 2.4%; Score 11; DB 2; Length 855;
 Best Local Similarity 100.0%; Pred. No. 0.044;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3

A29651

KEX1 protein precursor - Yeast (*Saccharomyces cerevisiae*)
 N/Alternate names: protein G1237; protein YGL203C
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 21-Jul-2000
 C/Accession: A29651; S64221
 R/Mochowska, A.; Dignard, D.; Henning, D.; Thomas, D.Y.; Bussey, H.
 Cell 50, 573-584, 1987
 A/Title: Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like funct
 A/Reference number: A29651; MUID:87273520; PMID:3301004
 A/Accession: A29651
 A/Molecule type: DNA
 A/Residues: 1-729 <DMO>
 A/Cross-references: GB:M17231; NID:G171778; PIDN:AAA34717.1; PID:G171779
 R/Bjournson, A.J.; McReynolds, A.D.K.; Wright, L.F.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64218
 A/Accession: S64221
 A/Molecule type: DNA
 A/Residues: 1-729 <BJO>
 A/Cross-references: EMBL:E72725; NID:G1322835; PID:6243802; PID:G1322836; MIPS:YGL203C
 A/Experimental source: strain S288C
 C/Genetics:
 A/Map position: 7L
 A/Map position: 7L
 A/Cross-references: SGD:S0003171; MIPS:YGL203C
 C/Keywords: Golgi apparatus; transmembrane protein
 F/1-22/Domain: signal sequence #status predicted <SIG>
 F/23-729/Product: KEX1 protein #status predicted <MAT>
 F/619-635/Domain: transmembrane #status predicted <TMM>

Query Match 2.2%; Score 10; DB 2; Length 729;
 Best Local Similarity 100.0%; Pred. No. 0.36;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDDDD 443
 DB 564 SEGDDDDDD 573

RESULT 4

S48244
 NM22 protein - Yeast (*Saccharomyces cerevisiae*)
 N/Alternate names: IFP1 protein; protein YHR077C
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 28-Oct-1994 #sequence_revision 10-Feb-1995 #text_change 23-Mar-2001
 C/Accession: S48244; S46815; S64648; S64738
 R/He, F.; Jacobson, A.
 submitted to the EMBL Data Library, September 1994
 A/Description: Identification of a novel component of the nonsense-mediated mRNA decay p
 A/Reference number: S48244
 A/Accession: S48244
 A/Molecule type: DNA
 A/Residues: 1-1089 <HEP>
 A/Cross-references: EMBL:U14974; NID:G555938; PID:G555939
 R/Favell, T.
 submitted to the EMBL Data Library, June 1994
 A/Description: The sequence of S. cerevisiae cosmid 9205.
 A/Reference number: S46795
 A/Accession: S46815
 A/Molecule type: DNA
 A/Residues: 'MYOQ', 3-1089 <FAV>
 A/Cross-references: EMBL:U10556; NID:G500825; PID:G500826; MIPS:YHR077C
 R/Lee, S.I.; Umen, U.G.; Varmus, H.E.
 Proc. Natl. Acad. Sci. U.S.A. 92, 6587-6591, 1995
 A/Title: A genetic screen identifies cellular factors involved in retroviral -1 frameshif
 A/Reference number: S64648; MUID:95327692; PMID:7604038
 A/Accession: S64648
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-1089 <LEE>
 A/Cross-references: EMBL:U28158
 R/Varmus, H.E.
 submitted to the EMBL Data Library, May 1995

A/Reference number: S64738
 A/Accession: S64738
 A/Molecule type: DNA
 A/Residues: 'MYOQ', 3-1089 <VAR>
 A/Cross-references: EMBL:U08158; NID:G967212; PID:G967213
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 24-Jan-1992 #sequence revision 24-Jan-1992 #text change 23-Mar-2001
 A/Title: The yeast MCK1 gene encodes a protein kinase homolog that activates early meiotic
 A/Reference number: A39622; MUID:91184610; PMID:2010083
 A/Accession: A39622
 A/Molecule type: DNA
 A/Residues: 1-375 <NEI>
 A/Cross-references: GB:X55054; NID:G3907; PIDN:CAA3895.1; PID:G3908
 R/Shero, J.H.; Hietter, P.
 Genes Dev. 5, 549-560, 1991
 A/Title: A suppressor of a centromere DNA mutation encodes a putative protein kinase (MCK1)
 A/Reference number: A39623; MUID:91184611; PMID:2010084
 A/Accession: A39623
 A/Molecule type: DNA
 A/Residues: 1-375 <SHE>
 A/Cross-references: GB:M55984; NID:G171909; PIDN:AAA34764.1; PID:G171910
 R/Dalley, D.; Schieven, G.L.; Lim, M.Y.; Marguardt, H.; Gilmore, T.; Thorne, J.; Martin
 Mol. Cell. Biol. 10, 6244-6256, 1990

Query Match 2.2%; Score 10; DB 2; Length 1089;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
 DB 902 DDDDDDDGEE 911

RESULT 5

hypothetical protein 11L - Molnuscum contagiosum virus 1
 N/Alternate names: MC011L
 C/Species: Molnuscum contagiosum virus 1
 C/Date: 05-Nov-1999 #sequence revision 05-Nov-1999 #text change 11-May-2000
 A/Accession: T30613
 R/Senevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A/Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
 A/Reference number: Z20876; MUID:9635459; PMID:8670425
 A/Accession: T30613
 A/Status: Preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-297 <SEN>
 A/Cross-references: EMBL:U60315; PIDN:AAC55139.1
 C/Genetics:
 A/Note: MC011L

Query Match 2.0%; Score 9; DB 2; Length 297;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDG 444
 DB 206 GDDDDDDG 214

RESULT 6

protein kinase MCK1 (EC 2.7.1.-) - yeast (*Saccharomyces cerevisiae*)
 N/Alternate names: protein kinase YPK1; protein N0392; protein YNL307C
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 24-Jan-1992 #sequence revision 24-Jan-1992 #text change 23-Mar-2001
 A/Accession: A39622; MUID:91184610; PMID:2010083
 A/Reference number: A39622; MUID:91184610; PMID:2010083
 A/Accession: A39622
 A/Molecule type: DNA
 A/Residues: 1-375 <NEI>
 A/Cross-references: GB:X55054; NID:G3907; PIDN:CAA3895.1; PID:G3908
 R/Shero, J.H.; Hietter, P.
 Genes Dev. 5, 549-560, 1991
 A/Title: A suppressor of a centromere DNA mutation encodes a putative protein kinase (MCK1)
 A/Reference number: A39623; MUID:91184611; PMID:2010084
 A/Accession: A39623
 A/Molecule type: DNA
 A/Residues: 1-375 <SHE>
 A/Cross-references: GB:M55984; NID:G171909; PIDN:AAA34764.1; PID:G171910
 R/Dalley, D.; Schieven, G.L.; Lim, M.Y.; Marguardt, H.; Gilmore, T.; Thorne, J.; Martin
 Mol. Cell. Biol. 10, 6244-6256, 1990

A/Title: Novel yeast protein kinase (YPK1 gene product) is a 40-kilodalton phosphotyrosyl
 A/Reference number: A36354; MUID:91061730; PMID:1701015
 A/Accession: A36354
 A/Molecule type: DNA
 A/Residues: 1-375 <DAI>
 A/Cross-references: GB:M55984; NID:G171909; PIDN:AAA34764.1; PID:G171910
 A/Accession: B36354
 A/Molecule type: protein
 A/Residues: 41-58; 70-94; 96-108; 134-139; 182-188; 238-254; 262-306 <DA2>
 R/Nicaud, J.C.
 Submitted to the EMBL Data Library, January 1995
 A/Description: Sequence analysis of a 13.9 kb fragment of yeast chromosome XIV identifies si

A/Reference number: S51285
 A/Accession: S51304
 A/Molecule type: DNA
 A/Residues: 1-375 <NIC>
 A/Cross-references: EMBL:Z46259; NID:G633655; PIDN:CAA63388.1; PID:G633675
 R/Mafteh, M.; Nicaud, J.M.; Levesque, H.; Galliard, C.
 Yeast 11, 1077-1085, 1995
 A/Title: Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV identifies si

A/Reference number: S59562; MUID:96076632; PMID:7502583
 A/Accession: S59573
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-375 <MAF>
 A/Cross-references: EMBL:Z46259; NID:G633655; PIDN:CAA63388.1; PID:G633675
 A/Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1995
 R/Mafteh, M.; Nicaud, J.M.; Levesque, H.; Galliard, C.
 Submitted to the Protein Sequence Database, April 1996
 A/Reference number: S63287
 A/Accession: S63288
 A/Molecule type: DNA
 A/Residues: 1-375 <MAF>
 A/Cross-references: EMBL:Z71583; NID:G1302406; PIDN:CAA96236.1; PID:G1302407; MIPS:YNL307C
 A/Experimental source: strain S288C
 R/Maurer, C.T.C.; Urbanus, J.H.M.; Plant, R.J.
 Submitted to the Protein Sequence Database, April 1996
 A/Reference number: S63286
 A/Accession: S63283
 A/Molecule type: DNA
 A/Residues: 1-17 <MAU>
 A/Cross-references: EMBL:Z71583; MIPS:YNL307C
 A/Experimental source: strain S288C
 C/Genetics:
 A/Note: SCD:MCK1; YPK1
 A/Cross-references: SCD:S0005251; MIPS:YNL307C
 A/Map position: 14L
 C/Function:
 A/Description: phosphotransferase
 C/Superfamily: kinase-related transforming protein; protein kinase homolog
 C/Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein ki
 F/33-302/Domain: protein kinase homolog <KIN>
 F/41-49/Region: protein kinase ATP-binding motif

Query Match 2.0%; Score 9; DB 2; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GVLIKIDFG 153
 DB 177 GVLIKIDFG 185

RESULT 7

protein kinase, 48K - *Chlamydomonas reinhardtii*
 C/Species: *Chlamydomonas reinhardtii*
 C/Date: 21-May-1999 #sequence revision 21-May-1999 #text change 08-Oct-1999
 A/Accession: T08069
 R/Kurvari, V.; Zhang, Y.; Luo, Y.; Snell, W.J.
 Proc. Natl. Acad. Sci. U.S.A. 93, 39-43, 1996
 A/Title: Molecular cloning of a protein kinase whose phosphorylation is regulated by gene
 A/Reference number: Z16331; MUID:96133873; PMID:8552645

A/Accession: T08069
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-408 <RUR>
A/Cross-references: EMBL:U36196; NID:G1019886; PIDN:AAA96956.1; PID:G1019887
A/Experimental source: strain 21gr(+)
C/Function:
A/Description: may be involved in a signalling pathway in fertilization
C/Keywords: autophosphorylation

Query Match 2.0%; Score 9; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDDDDDD 443
DB 393 EGDDDDDD 401

RESULT 8

JC4337
N:Activin receptor like kinase-1 homolog - mouse
N/Alternate names: activin type IB receptor
C/Species: Mus musculus (house mouse)
C/Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jul-2000
C/Accession: JC4337; I49478
R/Hub, X.; Robinson, C.E.; Pong, H.W.; Crabtree, J.S.; Rodriguez, B.R.; Roe, B.A.; Gimble
Biochem. Biophys. Res. Commun. 216, 78-83, 1995
A/Title: Cloning and characterization of the murine activin receptor like kinase-1 (ALK-
A/Reference number: JC4337; MUID:96067535; PMID:7488127
A/Accession: JC4337
A/Molecule type: mRNA

A/Residues: 1-502 <WUX>
A/Cross-references: GB:I48015; NID:G1408066; PIDN:AAB03642.1; PID:G1020393
A/Experimental source: lung
R/Dewulf, N.; Verschuuren, K.; Jonny, O.; Moren, A.; Grimsby, S.; Vande Spieghele, K.; M
Endocrinology 136, 2652-2663, 1995
A/Title: Distinct spatial and temporal expression patterns of two type I receptors for h
A/Reference number: I48241; MUID:95269711; PMID:7750469
A/Accession: I49478
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-502 <RES>
A/Cross-references: GB:I48015; NID:G1408066; PIDN:AAB03642.1; PID:G1020393
C/Comment: This protein belongs to transforming growth factor beta family, and it is a c
C/Genetics:

A/Gene: ALK-1
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; glycoprotein; receptor; transmembrane protein
F:118-139/Domain: transmembrane #status predicted <TMM>
F:199-495/Domain: protein kinase homology <KIN>
F:207-215/Region: protein kinase ATP-binding motif
F:97/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.0%; Score 9; DB 2; Length 502;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
DB 327 HRDLKSRNV 335

RESULT 9

ALK-1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: I48241
R/Dewulf, N.; Verschuuren, K.; Jonny, O.; Moren, A.; Grimsby, S.; Vande Spieghele, K.; M
Endocrinology 136, 2652-2663, 1995
A/Title: Distinct spatial and temporal expression patterns of two type I receptors for h
A/Reference number: I48241; MUID:95269711; PMID:7750469

A/Accession: I48241
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-502 <RES>
A/Cross-references: EMBL:U31664; NID:G840814; PIDN:CAA83484.1; PID:G840815
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP
F:199-495/Domain: protein kinase homology <KIN>
F:207-215/Region: protein kinase ATP-binding motif

Query Match 2.0%; Score 9; DB 2; Length 502;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
DB 327 HRDLKSRNV 335

RESULT 10

S64951
N:Hypothetical protein YIR114c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein L2941
C/Species: Saccharomyces cerevisiae
C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
C/Accession: S64951; S69401
R/Vernaseil, P.; Voel, M.; Volckaert, G.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S64943
A/Accession: S64951
A/Molecule type: DNA

A/Residues: 1-764 <VER>
A/Cross-references: EMBL:Z73286; NID:G1360509; PIDN:CAA87681.1; PID:G245807; PID:G136051
A/Experimental source: strain S288C
R/Vernaseil, P.; Volckaert, G.
submitted to the EMBL Data Library, September 1995
A/Reference number: S69393
A/Accession: S69401
A/Molecule type: DNA
A/Residues: 1-764 <VEW>
A/Cross-references: EMBL:X89514; NID:G1297019; PIDN:CAA61692.1; PID:E198747; PID:G129702
C/Genetics:
A/Gene: SGD:EFR4
A/Cross-references: SGD:S0004104
A/Map position: 12K

Query Match 2.0%; Score 9; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 445
DB 628 DDDDDDDG 636

RESULT 11

T52415
N:polycomb protein EZH1 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C/Accession: T52415
R/Biodeau, P.; Luo, M.; Dennis, E.S.; Peacock, W.J.; Chaudhury, A.M.
submitted to the EMBL Data Library, October 1998
A/Reference number: Z26069
A/Accession: T52415
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-856 <BIL>
A/Cross-references: EMBL:AF100163; PIDN:AAD09108.1

Query Match 2.0%; Score 9; DB 2; Length 856;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
|||||
DB 17 DDDDDGEE 25

RESULT 12

T01503
hypothetical protein T10M13.3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Mar-1999
C/Accession: T01503
R/Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Gott
Martensen, R.; McCombie, W.
submitted to the EMBL Data Library, May 1997
A/Description: The sequence of the Arabidopsis thaliana T10M13 BAC.
A/Reference number: Z14346
A/Accession: T01503
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-898 <OH>
A/Cross-references: EMBL:AF001308; NID:G2104523; PID:G3912918
A/Experimental source: cultivar Columbia
C/Genetics:
A/Map position: 4S
A/Intons: 48/2; 83/1; 118/3; 207/2; 260/2; 288/3; 336/3; 352/3; 554/2; 603/3; 677/2; 72
A/Note: T10M13.3

Query Match 2.0%; Score 9; DB 2; Length 898;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
|||||
DB 83 DDDDDGEE 91

RESULT 13

T43220
insulin-like growth factor-1 receptor - common lancelot
N/Alternate names: insulin-like peptide receptor
C/Species: Branchiostoma lanceolatum (common lancelet)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-May-2000
C/Accession: T43220
R/Pashmforoush, M.; Chan, S.-J.; Steiner, D.F.
Mol. Endocrinol. 10, 857-866, 1996
A/Title: Structure and expression of the insulin-like peptide receptor from amphioxus.
A/Reference number: Z22346; MUID:96408719; PMID:8813726
A/Accession: T43220
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1363 <PAS>
A/Cross-references: EMBL:883394; NID:G1911771; PID:G1911772; PIDN:AA850848.1
C/Superfamily: insulin receptor; protein kinase homology
C/Keywords: hormone receptor

Query Match 2.0%; Score 9; DB 2; Length 1363;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWEM 199
|||||
DB 1210 SYGVVLWEM 1218

RESULT 14

T30346
insulin receptor - yellow fever mosquito
C/Species: Aedes aegypti (yellow fever mosquito)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C/Accession: T30346
R/Graf, R.; Neuenchwander, S.; Brown, M.R.; Ackermann, U.
Insect Mol. Biol. 6, 151-163, 1996

A/Title: Insulin mediated secretion of ecdysteroids from mosquito ovaries and molecular

A/Reference number: Z20834

A/Accession: T30346

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1390 <GRA>

A/Cross-references: EMBL:U72939; NID:G1620749; PID:G1620750; PIDN:AA817094.1

C/Superfamily: insulin receptor; protein kinase homology

Query Match 2.0%; Score 9; DB 2; Length 1390;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWEM 199
|||||
DB 1224 SYGVVLWEM 1232

RESULT 15

T18410
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) II - malaria parasite
C/Species: Plasmodium falciparum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C/Accession: T18410
R/Plores, M.V.C.; O'Sullivan, M.J.; Stewart, T.S.
submitted to the EMBL Data Library, March 1997
A/Description: Characterisation of the carbamoyl phosphate synthetase gene from Plasmodi
A/Reference number: Z18931
A/Accession: T18410
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2391 <FLD>
A/Cross-references: EMBL:U32150; NID:G476023; PID:G476024; PIDN:AAA29522.1
C/Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Bac
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
F/96-2367/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>

Query Match 2.0%; Score 9; DB 2; Length 2391;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 445
|||||
DB 1786 DDDDDGEE 1794

RESULT 16

I38220
protein-serine/threonine kinase - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 24-Sep-1999
C/Accession: I38220; S37422
R/Schultz, S.J.; Ni99, E.A.
Cell Growth Differ. 4, 821-830, 1993
A/Title: Identification of 21 novel human protein kinases, including 3 members of a fami
A/Reference number: I38221; MUID:94100173; PMID:8274451
A/Accession: I38220
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-101 <RES>
A/Cross-references: EMBL:Z25430; NID:G405738; PIDN:CA80917.1; PID:G405739
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F/1-101/Domain: protein kinase homology (fragment) <KIN>

Query Match 1.8%; Score 8; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
|||||
DB 92 WMAPEVIQ 99

RESULT 17

138212 protein-serine/threonine kinase - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 24-Sep-1999
 C/Accession: 138212; S37414
 R/Schultz, S.J.; Nigg, E.A.
 Cell Growth Differ. 4, 821-830, 1993
 A/Title: Identification of 21 novel human protein kinases, including 3 members of a family
 A/Reference number: 138211; MUID:94100173; PMID:8274451
 A/Accession: 138212
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-108 <RBS>
 A/Cross-references: EMBL:Z25422; NID:g405722; PID:CAA80909.1; PID:g405723
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 F/1-108/Domain: protein kinase homology (fragment) <KIN>

Query Match 1.8%; Score 8; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred.No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
 DB 92 WMAPEVIQ 99

RESULT 18

S09814 hypothetical protein U51 - human cytomegalovirus (strain AD169)
 C/Species: human cytomegalovirus, human herpesvirus 5
 A/Note: host Homo sapiens (man)
 C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
 C/Accession: S09814
 R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
 M.; Barrett, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
 A/Reference number: S09749; MUID:90269039; PMID:2161319
 A/Accession: S09814
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-157 <CHS>
 A/Cross-references: EMBL:X17403; NID:g59591; PID:CAA35410.1; PID:g1780829
 A/Note: this sequence was submitted to the EMBL Data Library, December 1989
 C/Superfamily: varicella-zoster virus gene 25 protein

Query Match 1.8%; Score 8; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred.No. 8.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 DDDGEEB 447
 DB 12 DDDGEEB 19

RESULT 19

T05710 2S albumin precursor - soybean
 N/Alternate names: aspartic acid-rich peptide
 C/Species: Glycine max (soybean)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: T05710; A28485
 R/Wang, J.; Pichersky, E.
 Plant Physiol. 114, 1567, 1997
 A/Title: Nucleotide Sequence of S-Adenosyl-L-Methionine:caffeic Acid 3-O-Methyltransferase
 A/Reference number: Z15424
 A/Accession: T05710
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-158 <MAN>

A/Cross-references: EMBL:AF005030; NID:g2305019; PID:g2305020
 A/Experimental source: cultivar Hodgson 78; cotyledon
 R/odani, S.; Koide, T.; Ono, T.
 J. Biol. Chem. 262, 10502-10505, 1987
 A/Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(L-
 A/Reference number: A28485; MUID:87280104; PMID:3611081
 A/Accession: A28485
 A/Molecule type: protein

A/Residues: 22-64 <ODA>
 C/Superfamily: soybean 2S albumin
 C/Keywords: glycoprotein
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-64/Product: 2S albumin small chain #status predicted <MAT>
 F/54-56/Region: cell attachment (R-G-D) motif
 F/59/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 158;
 Best Local Similarity 100.0%; Pred.No. 8.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 DB 55 GDDDDDD 62

RESULT 20

B83130 hypothetical protein PA4129 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Dec-2002
 C/Accession: B83130
 R/Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri-
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Laidig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
 A/Reference number: A82950; MUID:20437337; PMID:10964043
 A/Accession: B83130
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-162 <STO>
 A/Cross-references: GB:AE004829; GB:AE004091; NID:g5950327; PID:AA07516.1; GSPDB:GN0013
 A/Experimental source: strain PA01
 C/Superfamily: uncharacterized conserved protein

Query Match 1.8%; Score 8; DB 2; Length 162;
 Best Local Similarity 100.0%; Pred.No. 8.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PLLPLAA 336
 DB 34 PLLPLAA 41

RESULT 21

A96520 hypothetical protein T2J15.12 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: A96520
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiz, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
 ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: A96520
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-198 <STO>
A/Cross-references: GB:AE005173; NID:g10645475; PIDN:AA621589.1; GSPDB:GN00141
C/Genetics:
A/Gene: T2J15.12
A/Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDD 441
DB 47 SEGDDDD 54

RESULT 22

D69413

conserved hypothetical protein AF1309 - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: D69413

R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Ouellette, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997

A/Authors: Uetrecht, T.; Cotton, M.D.; Springs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea

A/Reference number: A69250; PMID:9809343; PMID:9389475

A/Accession: D69413

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-211 <KLE>

A/Cross-references: GB:AE001013; GB:AE000782; NID:g2669336; PIDN:AA89938.1; PID:g264927

Query Match 1.8%; Score 8; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PLULPLAA 336
DB 174 PLULPLAA 181

RESULT 23

POBP82

antiterminal Q - phage 82

N/Alternate names: regulatory protein Q

C/Species: phage 82

C/Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 23-Jul-1999

C/Accession: A29791; S66585

R/Goliger, J.A.; Roberts, J.W.

J. Biol. Chem. 262, 11721-11725, 1987

A/Title: Bacteriophage 82 gene Q and Q protein: sequence, overproduction, and activity

A/Reference number: A29791; PMID:87308148; PMID:3624233

A/Accession: A29791

A/Molecule type: DNA

A/Residues: 1-229 <GOL>

A/Cross-references: GB:J02803; NID:g215364; PIDN:AAA3229.1; PID:g215365

A/Note: the authors translated the codon ACT for residue 188 as Phe and GAA for residue

R/Nahid, A.A.; Sharples, G.J.; Mandal, T.N.; Lloyd, R.G.

J. Mol. Biol. 257, 561-573, 1996

A/Title: Holliday junction resolvases encoded by homologous rusa genes in Escherichia coli

A/Reference number: S66579; PMID:96196428; PMID:8648624

A/Accession: S66585

A/Molecule type: DNA

A/Residues: 1-229 <MAH>

A/Cross-references: EMBL:X92588; NID:g1051111; PIDN:CMA6332.1; PID:g1051118

C/Genetics:

A/Gene: Q

C/Superfamily: phage 82 regulatory protein Q
C/Keywords: DNA binding; late protein; transcription regulation

Query Match 1.8%; Score 8; DB 1; Length 229;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297
DB 142 TLERLKL 149

RESULT 24

D90831

antiterminal protein [imported] - Escherichia coli (strain O157:H7, substrain RMD 05C)

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: D90831

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

Gawara, N.; Yasunaga, T.; Kuhara, S.; Shida, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A/Reference number: A99629; PMID:21156231; PMID:11258796

A/Accession: D90831

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-229 <HAY>

A/Cross-references: GB:BA000007; PIDN:BA835043.1; PID:g13361084; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RMD 0509952

C/Genetics:

A/Gene: EC81620

C/Superfamily: phage 82 regulatory protein Q

Query Match 1.8%; Score 8; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297
DB 142 TLERLKL 149

RESULT 25

G85688

probable antiterminal Q of prophage CP-933X Z1874 [imported] - Escherichia coli (strain

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C/Accession: G85688

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,

Iller, L.; Grobeck, E.J.; Davis, N.W.; Llin, A.; Dimantanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; PMID:21074935; PMID:11206551

A/Accession: G85688

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-229 <STO>

A/Cross-references: GB:AE005174; NID:g12514802; PIDN:AG55971.1; GSPDB:GN00145; UWGP:Z18

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: Z1874

C/Superfamily: phage 82 regulatory protein Q

Query Match 1.8%; Score 8; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297
DB 142 TLERLKL 149

RESULT 26

T47271
 hypothetical protein tmpr [imported] - Mycobacterium smegmatis insertion sequence IS1096
 C/Species: Mycobacterium smegmatis
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
 C/Accession: T47271
 R/Cirillo, J.D.; Barletta, R.G.; Bloom, B.R.; Jacobs, W.R.
 J. Bacteriol. 173, 7772-7780, 1991
 A/Title: A novel transposon trap for mycobacteria: isolation and characterization of IS1
 A/Reference number: Z24439; MUID:92078082; PMID:1660454
 A/Accession: T47271
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-237 <CTR>
 A/Cross-references: EMBL:M76495; NID:G150003; PIDN:AAA98488.1; PID:G150004
 A/Experimental source: insertion sequence IS1096
 C/Genetics:
 A/Gene: tmpr
 A/Mobile element: insertion sequence IS1096

Query Match 1.8%; Score 8; DB 2; Length 237;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 FDLSEGD 438
 |||||
 Db 119 FDLSEGD 126

RESULT 27
 T47611
 hypothetical protein T14E10.90 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 28-Jul-2000
 R/Cbernauer, B.; Ottenwaelter, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; W
 submitted to the Protein Sequence Database, March 2000
 A/Reference number: Z24463
 A/Accession: T47611
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-250 <OBE>
 A/Cross-references: EMBL:AJ138656
 A/Experimental source: cultivar Columbia; BAC clone T14E10
 C/Genetics:
 A/Map position: 3
 A/Insertions: 73/3; 129/3; 188/3
 A/Note: T14E10.90
 C/Superfamily: Arabidopsis thaliana hypothetical protein T14E10.90

Query Match 1.8%; Score 8; DB 2; Length 250;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 RLKRLERD 300
 |||||
 Db 54 RLKRLERD 61

RESULT 28
 G84645
 hypothetical protein At2G25220 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: G84645
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Neuen, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: G84645
 A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-260 <STO>
 A/Cross-references: GB:AE002093; NID:g4567255; PIDN:AA23669.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2G25220
 A/Map position: 2

Query Match 1.8%; Score 8; DB 2; Length 260;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHIRDKS 136
 |||||
 Db 125 VHIRDKS 132

RESULT 29
 A34599
 DNA-binding protein MCM1 - yeast (Saccharomyces cerevisiae)
 N/Alternate names: DNA-binding protein FUN80; DNA-binding protein PRPF/GRM; protein YM95;
 C/Species: Saccharomyces cerevisiae
 C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 31-Mar-2000
 C/Accession: A34599; S14888; S22855; S52892
 R/Ammerer, G.
 Genes Dev. 4, 299-312, 1990
 A/Title: Identification, purification, and cloning of a polypeptide (PRPF/GRM) that binds
 A/Reference number: A34599; MUID:90249735; PMID:2159934
 A/Accession: A34599
 A/Molecule type: DNA
 A/Residues: 1-286 <AMM>
 A/Cross-references: GB:X52453; NID:95266; PIDN:CAA36691.1; PID:95267
 J/Pasmore, S.; Maine, G.T.; Elble, R.; Christ, C.; Tye, B.K.
 J. Mol. Biol. 204, 593-606, 1988
 A/Title: Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary for
 A/Reference number: S14888; MUID:89141759; PMID:3066908
 A/Accession: S14888
 A/Molecule type: DNA
 A/Residues: 1-286 <PAS>
 A/Cross-references: EMBL:X14187; NID:g3909; PIDN:CAA32389.1; PID:g3910
 R/Dubois, E.; Bercy, J.; Descamps, F.; Messenguy, F.
 Gene 55, 265-275, 1987
 A/Title: Characterization of two new genes essential for vegetative growth in Saccharomyc
 A/Reference number: S22855; MUID:88030692; PMID:311883
 A/Accession: S22855
 A/Molecule type: DNA
 A/Residues: 1-8, 'S', '10-36, 'F', '38-155, 'AR' <DUB>
 A/Cross-references: EMBL:M17511; NID:g171524; PIDN:AAA3609.1; PID:g171525
 A/Note: the authors translated the codon TCT for residue 9 as Pro and TTC for residue 37
 R/Odell, C.; Bowman, S.
 submitted to the EMBL Data Library, February 1995
 A/Reference number: S52885
 A/Accession: S52885
 A/Molecule type: DNA
 A/Residues: 1-286 <ODE>
 A/Cross-references: EMBL:Z48502; NID:g695715; PIDN:CAA88409.1; PID:g695723; MIPS:YMR043w
 C/Genetics:
 A/Gene: SGD:MCM1
 A/Cross-references: SGD:S0004646; MIPS:YMR043w
 A/Map position: 13R
 C/Superfamily: transcription factor sqna; serum response factor DNA-binding domain homol
 C/Keywords: DNA binding; nucleus; transcription regulation
 F/17-72/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 1.8%; Score 8; DB 2; Length 286;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 Db 108 GDDDDDD 115

RESULT 30
 A37818

osteopontin precursor - mouse
 N:Alternate names: bone sialoprotein I; early T lymphocyte activation 1 protein; phospho
 C:Species: Mus musculus (house mouse)
 C>Date: 12-Feb-1993 #sequence_revision 27-Jun-1994 #ext_change 10-Sep-1999
 C/Accession: A37818; S04078; S12064; A33853; J10105; A60931; S11677
 J: Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
 J. Biol. Chem. 265, 14432-14438, 1990
 A>Title: The mouse osteopontin gene. Expression in monocytic lineages and complete nucle
 A:Reference number: A37818; MUID:90354433; PMID:2387863
 A:Accession: A37818
 A:Molecule type: DNA
 A:Residues: 1-294 <MT>
 A/Cross-references: EMBL:X51834
 R: Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
 Nucleic Acids Res. 17, 3296, 1989
 A>Title: Nucleotide sequence of cDNA for mouse osteopontin-like protein.
 A:Reference number: S04078; MUID:89263742; PMID:2726465
 A:Accession: S04078
 A:Molecule type: mRNA
 A:Residues: 1-294 <MT>
 A/Cross-references: EMBL:X13986; NID:G53755; PIDN:CAA32165.1; PID:G53756
 R: Yamamoto, S.
 submitted to the EMBL Data Library, January 1990
 A:Reference number: S12064
 A:Accession: S12064
 A:Molecule type: DNA
 A:Residues: 1-121, 'P', 123-294 <YAM>
 A/Cross-references: EMBL:X51834; NID:G53520; PIDN:CAA36132.1; PID:G297546
 R: Craig, A.M.; Smith, J.H.; Denhardt, D.T.
 J. Biol. Chem. 264, 9682-9689, 1989
 A>Title: Osteopontin, a transformation-associated cell adhesion phosphoprotein, is induc
 A:Reference number: A33853; MUID:89255479; PMID:2722855
 A:Accession: A33853
 A:Molecule type: mRNA
 A:Residues: 1-98, 'G', 100-294 <CRA>
 A/Cross-references: GB:J04806; NID:G200157; PIDN:AA57265.1; PID:G200158
 R: Patarca, R.; Freeman, G.J.; Singh, R.P.; Wei, F.Y.; Duffee, T.; Blattner, F.; Regnier,
 J. Exp. Med. 170, 145-161, 1989
 A>Title: Structural and functional studies of the early T lymphocyte activation 1 (Eta-1
 on.
 A:Reference number: J10105; MUID:89310352; PMID:2787378
 A:Accession: J10105
 A:Molecule type: mRNA
 A:Residues: 1-42, 'P', 44-294 <PAT>
 A/Cross-references: EMBL:X16151; NID:G50863; PIDN:CAA34276.1; PID:G50864
 R: Singh, R.P.; Patarca, R.; Schwartz, J.; Singh, P.; Cantor, H.
 J. Exp. Med. 171, 1931-1942, 1990
 A>Title: Definition of a specific interaction between the early T lymphocyte activation
 A:Reference number: A60931; MUID:90278349; PMID:2351930
 A:Accession: A60931
 A:Molecule type: protein
 A:Residues: 158-176 <SIN>
 C/Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,
 C/Genetics:
 A:Gene: Era-1
 A:Map position: 5
 A:Insertions: 18/3; 30/3; 57/3; 71/3; 165/3
 C/Superfamily: osteopontin
 C/Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-24/Product: osteopontin #status predicted <MAT>
 F:85-96/Region: aspartic acid-rich
 F:144-146/Region: cell attachment (R-G-D) motif
 F:78/Binding site: carbohydrate (asn) (covalent) #status predicted
 Query Match 1.8%; Score 8; DB 1; Length 294;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 31
 T52117
 zinc finger protein [imported] - Arabidopsis thaliana (fragment)
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #ext_change 20-Oct-2000
 C/Accession: T52117
 R: Hildeau, P.
 submitted to the EMBL Data Library, October 1997
 A:Reference number: Z25963
 A:Accession: T52117
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-298 <BL>
 A/Cross-references: EMBL:AF030304; PIDN:AAC09174.1
 Query Match 1.8%; Score 8; DB 2; Length 298;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
 DB 25 DDDDDDDG 32

RESULT 32
 JCS811
 osteopontin - rat
 C/Species: Rattus norvegicus (Norway rat)
 C>Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #ext_change 07-May-1999
 C/Accession: JCS811
 R: Iasa, M.; Chang, P.L.; Prince, C.W.; Pina, L.A.
 Biochem. Biophys. Res. Commun. 240, 602-605, 1997
 A>Title: Phosphorylation of osteopontin by Golgi apparatus casein kinase.
 A:Reference number: JCS811; MUID:98063283; PMID:9398611
 A:Accession: JCS811
 A:Molecule type: protein
 A:Residues: 1-301 <LAS>
 A:Experimental source: brain
 C/Comment: This protein is involved in the initiation of the bone calcification process,
 uses, early resistance to bacterial infection and binding of tumor cells at secondary site
 C/Superfamily: osteopontin
 Query Match 1.8%; Score 8; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
 DB 72 DDDDDDDG 79

RESULT 33
 A25917
 osteopontin precursor - rat
 N:Alternate names: bone sialoprotein I; phosphoprotein I, secreted; tumor-secreted phospho
 C/Species: Rattus norvegicus (Norway rat)
 C>Date: 18-Dec-1987 #sequence_revision 27-Jun-1994 #ext_change 10-Sep-1999
 C/Accession: A25917; A45132; B45132; S28772; S04506; A45925
 R: Oldberg, A.; Franzen, A.; Heinegard, D.
 Proc. Natl. Acad. Sci. U.S.A. 83, 8819-8823, 1986
 A>Title: Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveal
 A:Reference number: A25917; MUID:87067405; PMID:3024151
 A:Accession: A25917
 A:Molecule type: mRNA
 A:Residues: 1-317 <OLD>
 A/Cross-references: GB:M14656; NID:G205859; PIDN:AA41762.1; PID:G205860
 R: Singh, K.; Mukherjee, A.B.; De Vouge, M.W.; Mukherjee, B.B.
 U. Biol. Chem. 267, 23847-23851, 1992
 A>Title: Differential processing of osteopontin transcripts in rat kidney- and osteoblast
 A:Reference number: A45132; MUID:93054745; PMID:1429723
 A:Accession: A45132
 A:Molecule type: protein

A/Residues: 36-51 <SIN1>
 A/Experimental source: kidney
 A/Note: sequence extracted from NCBI backbone (NCBIP:118869)
 A/Accession: B45132
 A/Molecule type: protein
 A/Residues: 272-282 <SIN2>
 A/Note: sequence extracted from NCBI backbone (NCBIP:118871)
 R/Prince, C.W.; Oosawa, T.; Butler, W.T.; Tomana, M.; Bhown, A.S.; Bhown, M.; Schreier, J. Biol. Chem. 282, 2900-2907, 1997
 A/Title: Isolation, characterization, and biosynthesis of a phosphorylated glycoprotein
 A/Reference number: S28772; MUID:87137549; PMID:3469201
 A/Accession: S28772
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 17-26 'X' <PRI>
 R/Senger, D.R.; Ferruzzi, C.A.; Papadopoulos, A.; Tenen, D.G.
 Biochim. Biophys. Acta 996, 43-48, 1989
 A/Title: Purification of a human milk protein closely similar to tumor-secreted phosphoglycoprotein
 A/Reference number: S04505; MUID:89287357; PMID:2736258
 A/Accession: S04506
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 17-26;155-167 <SEN>
 R/Senger, D.R.; Ferruzzi, C.A.; Gracey, C.F.; Papadopoulos, A.; Tenen, D.G.
 Cancer Res. 48, 5770-5774, 1988
 A/Title: Secreted phosphoproteins associated with neoplastic transformation: close homology with the osteopontin gene
 A/Reference number: A45925; MUID:89002730; PMID:3167835
 A/Accession: A45925
 A/Molecule type: protein
 A/Residues: 17-25 <SE2>
 C/Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid, C/Superfamily: osteopontin
 C/Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-317/Product: osteopontin #status predicted <MAT>
 F:66-96/Region: aspartic acid-rich
 F:144-146/Region: cell attachment (R-G-D) motif
 F:79/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 317;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
 Db 88 DDDDDDDG 95

RESULT 34
 E69026
 conserved hypothetical protein MTH1196 - Methanobacterium thermoautotrophicum (strain Delta H)
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 26-May-2000
 A/Accession: E69026
 R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Alredge, T.; Liu, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Kt. S.; Church, G.M.; Daniels, C.J.; Mao, V.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7151, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional genome
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: E69026
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-318 <MTH>
 A/Cross-references: GB:AE00088; GB:AE00066; NID:G2622304; PIDN:AA85685.1; PID:G262230
 A/Experimental source: strain Delta H
 C/Genetics:
 A/Genes: MTH1196
 A/Start codon: TTG
 C/Superfamily: conserved hypothetical protein MTH1196

Query Match 1.8%; Score 8; DB 2; Length 318;
 Best Local Similarity 100.0%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 VKKLKIE 51
 Db 28 VKKLKIE 35

RESULT 35
 T16747
 hypothetical protein R13F6.7 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
 A/Accession: T16747
 R/Miller, N.
 submitted to the EMBL Data Library, April 1994
 A/Description: The sequence of C. elegans cosmid R13F6.
 A/Reference number: Z18570
 A/Accession: T16747
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-328 <MIL>
 A/Cross-references: EMBL:U0046; NID:G470358; PID:G470364; PIDN:AA47047.1; GSPDB:GN00021
 A/Experimental source: strain Bristol NZ; Clone R13F6
 C/Genetics:
 A/Genes: CESP:R13F6.7
 A/Map position: 3
 A/Introns: 20/3; 160/3; 222/2; 286/2
 C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 hc

Query Match 1.8%; Score 8; DB 2; Length 328;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 EKEAEILS 58
 Db 90 EKEAEILS 97

RESULT 36
 R5HSL0
 ribosomal protein L10 [similarity] - Halobacterium salinarum
 N/Alternate names: ribosomal protein P0
 C/Species: Halobacterium salinarum
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 21-Jul-2000
 A/Accession: S01315; S04120; S11587
 R/Iton, T.
 Eur. J. Biochem. 176, 297-303, 1988
 A/Title: Complete nucleotide sequence of the ribosomal 'A' protein operon from the archaeon Halobacterium salinarum
 A/Reference number: S01314; MUID:88329082; PMID:2458258
 A/Accession: S01315
 A/Molecule type: DNA
 A/Residues: 1-352 <ITO>
 A/Cross-references: EMBL:X13008; NID:G43532; PIDN:CAA3131.1; PID:G43534
 A/Experimental source: strain S9
 A/Note: the source is designated as Halobacterium halobium
 R/Shimizu, L.C.; Dennis, P.P.
 EMBO J. 8, 1225-1235, 1989
 A/Title: Characterization of the L11, L1, L10 and L12 equivalent ribosomal protein gene
 A/Reference number: S04116; MUID:89305527; PMID:2743281
 A/Accession: S04120
 A/Molecule type: DNA
 A/Residues: 1-56; 'V', 60-352 <SHI>
 A/Cross-references: EMBL:X15078; NID:G43449; PIDN:CAA33180.1; PID:G43454
 A/Note: the source is designated as Halobacterium cutitubrum
 C/Superfamily: rat acidic ribosomal protein P0
 C/Keywords: protein biosynthesis; ribosome

Query Match 1.8%; Score 8; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
 Db 88 DDDDDDDG 95

Db 333 DDDDDDDG 340

RESULT 37

G84266

50S ribosomal protein L10P [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: G84266

R:Jing, W.V.; Kennedy, S.P.; Mahaffas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.

R:Lehman, B.; Keller, K.; Cruz, R.; Danon, M.J.; Hough, D.W.; Maddocks, D.G.; Jaid

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.; Li

A>Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; WUID:20504483; PMID:11016950

A:Accession: G84266

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-352 <STO>

A:Cross-references: GB:AE004437; NID:g10580651; PIDN:AA019499.1; GSPDB:GN00138

C:Genetics:

A:Gene: rpl10P

C:Superfamily: rat acidic ribosomal protein P0

Query Match 1.8%; Score 8; DB 2; Length 352;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444

Db 333 DDDDDDDG 340

RESULT 38

G96668

protein F1N19.7 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 01-Mar-2002

C:Accession: G96668

R:Theologis, A.; Eckert, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; WUID:21016719; PMID:11130712

A:Accession: G96668

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <STO>

A:Cross-references: GB:AE005173; NID:g6633811; PIDN:AAF19670.1; GSPDB:GN00141

C:Genetics:

A:Gene: F1N19.7

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein T4P9.90

Query Match 1.8%; Score 8; DB 2; Length 368;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444

Db 61 DDDDDDDG 68

RESULT 39

T33173

hypothetical protein C24G6.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33173

R:Greco, T.; Bradshaw, H.; Kepler, D.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid C24G6.

A:Reference number: 221298

A:Accession: T33173

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-374 <GRE>

A:Cross-references: EMBL:AF067936; PIDN:AA019208.1; GSPDB:GN00023; CESP:C24G6.5

A:Experimental source: strain Bristol N2; clone C24G6

C:Genetics:

A:Gene: CESP:C24G6.5

A:Map position: 5

A:introns: 83/1; 351/3

Query Match 1.8%; Score 8; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443

Db 189 GDDDDDD 196

RESULT 40

T04645

hypothetical protein F10N7.190 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Sep-1999

C:Accession: T04645

R:Devan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hebeisel, J.; Mewes, H.W.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15263

A:Accession: T04645

A:Molecule type: DNA

A:Residues: 1-379 <BEV>

A:Cross-references: EMBL:AL021636

A:Experimental source: cultivar Columbia; BAC clone F10N7

C:Genetics:

A:Map position: 4

A:introns: 48/1; 221/1; 252/3; 287/1; 326/3

A:Note: F10N7.190

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 1.8%; Score 8; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 VIHRLDKS 136

Db 233 VIHRLDKS 240

RESULT 41

T01451

protein kinase homolog F2401.13 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001

C:Accession: T01451

R:Shim, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con

eologis, A.; Eckert, J.R.

submitted to the EMBL Data Library, January 1998

A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.

A:Reference number: Z14211

A:Accession: T01451

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-390 <SHI>

A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781357; GSPDB:GN00059; ATSP:F2401.

A:Gene: ATSP:F2401.13

A:Map position: 1

A:introns: 149/3; 301/3

C:Superfamily: Kinase-related transforming protein, protein kinase homology

Query Match

1.8%; Score 8; DB 2; Length 390;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHIRDLKS 136

DB 208 VHIRDLKS 215

RESULT 42

D64366

hypothetical protein homolog MJ0532 - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C/Accession: D64366

R/Bull: C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.

erson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Huxst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese,

A>Title: Complete genome sequence of the mechanogenic archaeon, *Methanococcus jannaschii*

A:Reference number: A64300; MUID:96337999; PMID:868087

A:Accession: D64366

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-391 <BUL>

A:Cross-references: GB:067502; GB:L77117, NID:92826293; PIDN:AB98523.1; PID:g1591235; T

C:Genetics:

A:Map position: REV468785-467610

C:Superfamily: hypothetical protein MJ0532

Query Match

1.8%; Score 8; DB 2; Length 391;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVIADGV 146

DB 149 VVIADGV 156

RESULT 43

S52578

serine/threonine-specific protein kinase NPK15 (EC 2.7.1.-) - common tobacco

C:Species: *Nicotiana tabacum* (common tobacco)

C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999

C/Accession: S52578

R/Ito, Y.; Banno, H.; Moribe, T.; Hinata, K.; Machida, K.H.Y.

Mol. Gen. Genet. 245, 1-10, 1994

A>Title: NPK15, a tobacco protein-serine/threonine kinase with a single hydrophobic reg

A:Reference number: S52578; MUID:95147840; PMID:784531

A:Accession: S52578

A:Molecule type: mRNA

A:Residues: 1-422 <ITO>

A:Cross-references: EMBL:D11737, NID:g505145, PIDN:BA06538.1; PID:d1007109; PID:g505146

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:114-182/Domain: protein kinase homology <KIN>

F:122-130/Region: protein kinase ATP-binding motif

QY 129 VHIRDLKS 136

DB 235 VHIRDLKS 242

RESULT 44

T05676

hypothetical protein F20M13.40 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C/Accession: T05676

R:Bevan, M.; Weller, H.; Kutner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15420

A:Accession: T05676

A:Molecule type: DNA

A:Residues: 1-426 <BEV>

A:Cross-references: EMBL:AL035540

A:Experimental source: cultivar Columbia; BAC clone F20M13

A:Map position: 4

A:introns: 90/3; 149/2; 204/1; 350/3

A>Note: F20M13.40

Query Match

1.8%; Score 8; DB 2; Length 426;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443

DB 407 GDDDDDD 414

RESULT 45

B54843

nemo, form II - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 19-Dec-1997

C/Accession: B54843

R/Choi, K.W.; Benzer, S.

Cell 78, 125-136, 1994

A>Title: Rotation of photoreceptor clusters in the developing *Drosophila* eye requires the

A:Reference number: A54843; MUID:94306509; PMID:8032304

A:Accession: B54843

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-434 <CHO>

A:Cross-references: GB:U12009

C:Genetics:

A:Gene: nmo

A:Cross-references: FlyBase:FBgn0011817

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: alternative splicing; ATP

F:38-301/Domain: protein kinase homology <KIN>

F:46-54/Region: protein kinase ATP-binding motif

Query Match

1.8%; Score 8; DB 2; Length 434;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153

DB 179 VLKICDFG 186

RESULT 46

T51402

serine/threonine-specific protein kinase-like protein - *Arabidopsis thaliana*

N/Alternate names: protein F14F8_110

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000

C/Accession: T51402

R:Atto, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51402

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <SAT>
A:Cross-references: EMBL:AL391144
A:Experimental source: cultivar Columbia; BAC clone F14F8
C:Genetics:
A:Map position: 5
A:Introns: 107/2; 158/3; 202/1; 254/3; 322/3
A:Note: F14F8.110
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 1.8%; Score 8; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLKS 136
DB 235 VHRDLKS 242

RESULT 47
T31734
hypothetical protein T05C3.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C:Accession: T31734
R:Blanchard, M.; Bradshaw, H.; Kramer, J.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid T05C3.
A:Reference number: Z21076
A:Accession: T31734
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-439 <BLA>
A:Cross-references: EMBL:AF014428; PIDN:AAB65361.1; GSPDB:GN00023; CESP:T05C3.5
A:Experimental source: strain Bristol N2; clone T05C3
C:Genetics:
A:Gene: CESP:T05C3.5
A:Map position: 5
A:Introns: 267/3; 298/2; 377/3
C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

Query Match 1.8%; Score 8; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 105 GDDDDDD 112

RESULT 48
T39232
probable serine threonine protein kinase - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39232
R:Churcher, C.M.; Gentsles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21837
A:Accession: T39232
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-471 <CHU>
A:Cross-references: EMBL:Z98763; PIDN:CA811493.1; GSPDB:GN00066; SPDB:SPAC9G1.09
A:Experimental source: strain 972h-; cosmid c9G1
C:Genetics:
A:Gene: SPDB:SPAC9G1.09
A:Map position: 1
A:Introns: 39/3
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 1.8%; Score 8; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 171 WMAPEVIQ 178

RESULT 49
A54843
nemo, form I - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 24-Sep-1999
C:Accession: A54843
R:Choi, K.W.; Benzer, S.
Cell 76, 125-136, 1994
A:Title: Rotation of photoreceptor clusters in the developing *Drosophila* eye requires the
A:Reference number: A54843; MUID:94306509; PMID:8033204
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <CHO>
A:Cross-references: GB:U12009; NID:9515669; PIDN:AAA21124.1; PID:9532558
C:Genetics:
A:Gene: nmo
A:Cross-references: FlyBase:FBgn0011817
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: alternative splicing; ATP
F:38-301/Domain: protein kinase homology <KIN>
F:46-54/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKIDFG 153
DB 179 VLKIDFG 186

RESULT 50
S37845
transcription initiation factor IIE chain TFA1 - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YKL028w; transcription initiation factor a 66k chain
C:Species: *Saccharomyces cerevisiae*
A:Variety: strain S288C
C:Date: 23-Feb-1995 #sequence_revision 11-Aug-1995 #text_change 21-Jul-2000
C:Accession: S37845; A55465; C55465
R:Rieger, M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37832
A:Accession: S37845
A:Molecule type: DNA
A:Residues: 1-482 <RIE>
A:Cross-references: EMBL:Z28028; NID:9486026; PID:9486027; MIPS:YKL028w
R:Reaver, W.J.; Henry, N.L.; Bushnell, D.A.; Saye, M.H.; Brickner, J.H.; Gilardi, O.; K
J. Biol. Chem. 269, 27549-27553, 1994
A:Title: Yeast TFIIE. Cloning, expression, and homology to vertebrate proteins.
A:Reference number: A55465; MUID:95050500; PMID:7961670
A:Accession: A55465
A:Molecule type: DNA
A:Residues: 1-482 <FEA>
A:Cross-references: GB:U18825; NID:9607957; PIDN:AAA62665.1; PID:9607958
A:Accession: C55465
A:Molecule type: protein
A:Residues: 102-108; 206-220 <FE2>
C:Genetics:
A:Gene: SGD:TFA1
A:Cross-references: SGD:S0001511; MIPS:YKL028w
A:Map position: 11L
C:Keywords: nucleus; transcription initiation

Query Match 1.8%; Score 8; DB 2; Length 482;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 DB 464 GDDDDDD 471

RESULT 51

T49237
 hypothetical protein F7K15.80 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C/Accession: T49237

R/Obermaier, B.; Ottenwelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lem
 submitted to the Protein Sequence Database, April 2000

A/Reference number: Z25019

A/Accession: T49237

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-485 <OBE>

A/Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7K15.80

A/Experimental source: cultivar Columbia; BAC clone F7K15

C/Genetics:

A/Map position: 3

A/Introns: 292/3; 359/3; 425/1

Query Match 1.8%; Score 8; DB 2; Length 485;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 DB 25 DDDDDDDG 32

RESULT 52

D84860
 probable protein kinase [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: D84860

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: AB4420; MUID:20083487; PMID:10617197

A/Accession: D84860

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-494 <STO>

A/Cross-references: GB:AE002093; NID:G4512659; PIDN:AAD21713.1; GSPDB:GN00139

C/Genetics:

A/Map position: 2

Query Match 1.8%; Score 8; DB 2; Length 494;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVAVKKLL 48
 DB 207 EVAVKKLL 214

RESULT 53

T10558

hypothetical protein T12G13.150 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C/Accession: T10558

R/Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
 submitted to the Protein Sequence Database, June 1999

A/Reference number: Z16533

A/Accession: T10558

A/Molecule type: DNA

A/Residues: 1-504 <BEV>

A/Cross-references: EMBL:AL080252; GSPDB:GN00062; ATSP:T12G13.150

A/Experimental source: cultivar Columbia; BAC clone T12G13

C/Genetics:

A/Map position: 4

A/Introns: 54/2; 89/3; 192/2; 227/3; 334/3; 360/2; 402/1; 496/1

Query Match 1.8%; Score 8; DB 2; Length 504;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 DKEVAVKK 46
 DB 274 DKEVAVKK 281

RESULT 54

B96524

hypothetical protein F11A17.5 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C/Accession: B96524

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maiti, R.; Marzilli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A66141; MUID:21016719; PMID:11130712

A/Accession: B96524

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-513 <STO>

A/Cross-references: GB:AE005173; NID:G5733869; PIDN:AAD49757.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: Arabidopsis thaliana hypothetical protein F17J16.30

Query Match 1.8%; Score 8; DB 2; Length 513;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 DB 319 DDDDDDDG 326

RESULT 55

T47786

hypothetical protein F17J16.90 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C/Accession: T47786

R/D'Angelio, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000

A/Reference number: Z24476

A/Accession: T47786

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-526 <DAN>

A/Cross-references: EMBL:AL163527

A/Experimental source: cultivar Columbia; BAC clone F17J16

C/Genetics:

A/Map position: 3

A;introns: 67/3; 247/3; 289/2; 322/2
A;Note: F17J16.90

Query Match 1.8%; Score 8; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDD 442
|||||
DB 491 EGDGDDDD 498

RESULT 56

C83284
probable biotin-dependent carboxylase PA2888 [imported] - Pseudomonas aeruginosa (strain
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: C83284
R;Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bh
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Lam,
.; Lory, S.; Olson, M.V
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; PMID:20437337; PMID:10984043
A;Accession: C83284
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-538 <STO>
A;Cross-references: GB:AE004715; GB:AE004091; NID:G9948977; PIDN:AA606276.1; GSPDB:GN001
C;Genetics:
A;Gene: PA2888
C;Superfamily: proplyl-CoA carboxylase beta chain

Query Match 1.8%; Score 8; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 EGSNTPL 331
|||||
DB 364 EGSNTPL 371

RESULT 57

AI2238
glutamine-fructose-6-P-aminotransferase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AI2238
R;Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; PMID:21595285; PMID:11759840
A;Accession: AI2238
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-541 <RUR>
A;Cross-references: GB:BA000019; PIDN:BA75163.1; PID:G17132597; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: nodM
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

Query Match 1.8%; Score 8; DB 2; Length 541;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 EIEATLER 293
|||||
DB 375 EIEATLER 382

RESULT 58

T49839
Hypothetical protein B24H17.250 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49839
R;Schulte, U.; Altm, V.; Hohnselt, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49839
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-560 <SCH>

A;Cross-references: EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.250
A;Experimental source: BAC clone B24H17; strain OR74A
C;Genetics:
A;Gene: NCSP:B24H17.250
A;Map position: 6
A;introns: 72/1; 142/3; 270/1; 310/1; 380/3; 411/2; 446/2; 468/1; 485/3

Query Match 1.8%; Score 8; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||||
DB 385 DDDDDDDG 392

RESULT 59

JCS957
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C;Accession: JCS957
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A;Reference number: JCS955; PMID:98153801; PMID:9480845
A;Accession: JCS957
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-567 <SAK>
A;Cross-references: DDBJ:AB009358
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 567;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
|||||
DB 170 VLKICDFG 177

RESULT 60

JCS955
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: JCS955
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A;Reference number: JCS955; PMID:98153801; PMID:9480845
A;Accession: JCS955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-579 <SAK>
A;Cross-references: DDBJ:AB009356; NID:G2924623; PIDN:BA25025.1; PID:G2924624
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C/Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 579;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
 |||||
 DB 170 VLKICDFG 177

RESULT 61

T49206

phosphoinositide-specific phospholipase C-like protein - Arabidopsis thaliana

N/Alternate names: protein F27K19.120

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 28-Jul-2000

C/Accession: T49206

R/Benes, V.; Murnbach, B.; Drzonek, H.; Ansgorge, W.; Newes, H.W.; Rudd, S.; Lemcke, K.; submitted to the Protein Sequence Database, April 2000

A/Reference number: Z25014

A/Accession: T49206

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-584 <BEN>

A/Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.120

A/Experimental source: cultivar Columbia; BAC clone F27K19

C/Genetics:

A/Map position: 3

A/Intons: 100/3; 166/2; 211/3; 290/3; 371/2; 410/3; 458/3; 487/3

C/Species: Arabidopsis thaliana 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase

F/105-248/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain x hom

Query Match 1.8%; Score 8; DB 2; Length 584;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 DB 295 GDDDDDD 302

RESULT 62

S57594

hypothetical protein YMR27c - Yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YM959.09c

C/Species: Saccharomyces cerevisiae

C/Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Oct-1995

C/Accession: S57594

R/Skelton, J.; Churcher, C.M. submitted to the EMBL Data Library, June 1995

A/Reference number: S57587

A/Accession: S57594

A/Molecule type: DNA

A/Residues: 1-590 <SKC>

A/Cross-references: EMBL:Z49939; NID:g887599; PIDN:CAA90198.1; PID:g887608; MIPS:YMR227C

A/Experimental source: strain AB972

C/Genetics:

A/Map position: 13R

Query Match 1.8%; Score 8; DB 2; Length 590;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 |||||
 DB 413 DDDDDDDG 420

RESULT 63

JC5956

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human

C/Species: Homo sapiens (man)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C/Accession: JC5956

R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T. Biochem. Biophys. Res. Commun. 243, 545-549, 1998

A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind

A/Reference number: JC5955; M0ID:98153801; PMID:9480845

A/Accession: JC5956

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-606 <SAK>

A/Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BA25026.1; PID:g2924626

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: C96596

R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hulzar, L. Nature 408, 816-820, 2000

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; M0ID:21016719; PMID:11130712

A/Accession: C96596

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-684 <STO>

A/Cross-references: GB:AE005173; NID:g11094789; PIDN:AG29721.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 DB 162 GDDDDDD 169

RESULT 65

C96572

protein F12M16.4 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: C96572

R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hulzar, L. Nature 408, 816-820, 2000

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; M0ID:21016719; PMID:11130712

A/Accession: C96572

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-684 <STO>

A/Cross-references: GB:AE005173; NID:g11094789; PIDN:AG29721.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 DB 162 GDDDDDD 169

QY 437 DDDDDDDG 444
 |||||
 DB 413 DDDDDDDG 420

QY 437 DDDDDDDG 444
 |||||
 DB 413 DDDDDDDG 420

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96572
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-690 <STO>
 A:Cross-references: GB:AE005173; NID:g7769851; PIDN:AAF69529.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F12M16.4
 A:Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 690;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIQ 179
 DB 187 WMAPEVIQ 194

RESULT 66
 G02838
 enhancer-of-zeste homolog 2 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G02838
 R:Antonarakis, S.E.
 A:Submitted to the EMBL Data Library, February 1996
 A:Reference number: H01746
 A:Accession: G02838
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-746 <ANT>
 A:Cross-references: EMBL:X85653; NID:g1438063
 C:Genetics:
 A:Gene: GDB:EZH2; EZH1
 A:Cross-references: GDB:701613; OMIM:601573
 A:Map position: 21q22.2-21q22.2

Query Match 1.8%; Score 8; DB 2; Length 746;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
 DB 183 DDDDDDDG 190

RESULT 67
 T02852
 probable membrane protein L1439.4 [imported] - *Leishmania* major (strain Friedlin)
 C:Species: *Leishmania* major
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C:Accession: T02852; H81462
 R:Myler, P.J.
 A:Submitted to the EMBL Data Library, May 1998
 A:Description: The nucleotide sequence of *Leishmania* major Friedlin chromosome 1.
 A:Reference number: Z14740
 A:Accession: T02852
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-800 <MTL>
 A:Cross-references: EMBL:AF001274; NID:g3264850; PID:g2266920
 R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A>Title: *Leishmania* major Friedlin chromosome 1 has an unusual distribution of protein-c
 A:Reference number: A81455; MUID:99178987; PMID:10077609
 A:Accession: H81462
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-800 <EYL>
 A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AAAC4675.1; PID:g2266920; GSPDB:GN001
 A:Experimental source: strain MHOM/IL/81/Friedlin
 C:Genetics:
 A:Gene: L1439.4
 A:Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 800;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
 DB 21 DDDDDDDG 28

RESULT 68
 B96716
 probable serine/threonine kinase F23010.20 [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96716
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizart, L.
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B96716
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-836 <STO>
 A:Cross-references: GB:AE005173; NID:g7705100; PIDN:AAF67779.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F23010.20
 A:Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 836;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIQ 179
 DB 412 WMAPEVIQ 419

RESULT 69
 S48975
 hypothetical protein YHR131c - yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
 C:Accession: S48975
 R:Fullon, L.
 A:Submitted to the EMBL Data Library, June 1994
 A:Description: The sequence of *S. cerevisiae* cosmid 9315.
 A:Reference number: S48967
 A:Accession: S48975
 A:Molecule type: DNA
 A:Residues: 1-840 <FUL>
 A:Cross-references: EMBL:U10398; NID:g551328; PID:g500681; GSPDB:GN00008; MIPS:YHR131c
 C:Genetics:
 A:Gene: MIPS:YHR131c
 A:Cross-references: SGD:S0001173
 A:Map position: 8R

Query Match 1.8%; Score 8; DB 2; Length 840;
 Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 444
 |||||
 Db 829 DDDDDDD 836

RESULT 70

A53800
 mixed-lineage protein kinase (EC 2.7.1.1 -) 3 - human
 N:Alternate names: protein kinase PTK1; protein kinase SPRK
 C:Species: Homo sapiens (man)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A53800; 158395
 R:Galio, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
 J. Biol. Chem. 269, 15092-15100, 1994
 A>Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
 A:Reference number: A53800; MUID:94253068; PMID:8195146
 A:Accession: A53800

A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-847 <GAL>
 A:Cross-references: GB:U0747; NID:9464027; PIDN:AA19647.1; PID:9464028
 R:ing, Y.L.; Leung, I.W.; Heng, H.H.; Teui, L.C.; Lassam, N.J.
 Oncogene 9, 1745-1750, 1994
 A>Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain
 A:Reference number: 158395; MUID:94239754; PMID:8183572
 A:Accession: 158395
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-847 <RES>

A:Cross-references: GB:U32976; NID:9488295; PIDN:AA59859.1; PID:9488296
 C:Genetics:
 A:Gene: GDB:MLK3; PTK1; SPRK
 A:Cross-references: GDB:134755; OMIM:600050
 A:Map position: 11q13.1-11q13.3
 C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
 F:48-100/Domain: SH3 homology <SH2>
 F:115-383/Domain: protein kinase homology <KIN>
 F:123-131/Region: protein kinase ATP-binding motif
 F:403-424/Region: leucine zipper motif
 F:438-459/Region: leucine zipper motif
 F:468-482/Region: basic

Query Match 1.8%; Score 8; DB 1; Length 847;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLKS 136
 |||||
 Db 237 VHRDLKS 244

RESULT 71

T18861
 probable peptide-aspartate beta-dioxygenase (EC 1.14.11.16) - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T18861; T23533
 R:Swindburne, J.
 submitted to the EMBL Data Library, August 1996

A:Reference number: Z19032
 A:Accession: T18861
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-872 <WTL>
 A:Cross-references: EMBL:Z79596; NID:91323798; PIDN:CAB01859.1; GSPDB:GN00028; CESP:K09A
 R:Swindburne, J.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19753
 A:Accession: T23533

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-872 <WIL>
 A:Cross-references: EMBL:Z79601; PIDN:CAB01887.1; GSPDB:GN00028; CESP:K09A9.6
 A:Experimental source: clone K09A9

C:Genetics:
 A:Gene: CESP:K09A9.6
 A:Map position: X
 A:Intons: 18/1; 40/3; 74/1; 104/2; 139/1; 170/3; 385/3; 432/2; 464/3; 530/2; 609/3; 776/
 C:Keywords: oxidoreductase

Query Match 1.8%; Score 8; DB 2; Length 872;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 Db 288 GDDDDDD 295

RESULT 72

A56731
 Chromatin assembly factor I p150 chain - human
 C:Species: Homo sapiens (man)
 C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 21-Jul-2000
 C:Accession: A56731
 R:Nautman, P.D.; Kobayashi, R.; Kessler, N.; Stillman, B.
 Cell 81, 1105-1114, 1995
 A>Title: The p150 and p60 subunits of chromatin assembly factor I: a molecular link betwe
 A:Reference number: A56731; MUID:95323966; PMID:7600578
 A:Accession: A56731

A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-938 <KAU>
 A:Cross-references: GB:U20979; NID:9882257; PIDN:AA46736.1; PID:9882258

Query Match 1.8%; Score 8; DB 2; Length 938;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDD 441
 |||||
 Db 598 SEGDDDD 605

RESULT 73

B45082
 neurotrophic receptor rox2 precursor - human
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
 C:Accession: B45082
 R:Masiakowski, P.; Carroll, R.D.
 J. Biol. Chem. 267, 26181-26190, 1992
 A>Title: A novel family of cell surface receptors with tyrosine kinase-like domain.
 A:Reference number: A45082; MUID:93100347; PMID:1334494
 A:Accession: B45082

A:Molecule type: mRNA
 A:Residues: 1-943 <MAS>
 A:Cross-references: GB:M97639; NID:9337466; PIDN:AA460276.1; PID:9337467
 A>Note: sequence extracted from NCBI backbone (NCBIP:1120918)
 C:Genetics:
 A:Gene: GDB:NTRKR2

A:Cross-references: GDB:136454
 A:Map position: 6p21-6p21
 C:Superfamily: neurotrophic receptor for; immunoglobulin homology; kringle homology; prot
 C:Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyrosi
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-937/Domain: immunoglobulin homology <IM>
 F:76-137/Domain: immunoglobulin homology <IM>
 F:316-394/Domain: kringle homology <KR>
 F:412-428/Domain: transmembrane #status predicted <TM>
 F:471-753/Domain: protein kinase homology <KIN>

F:479-487/Region: protein kinase ATP-binding motif
F:70,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 2; Length 943;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWE 198

DB 677 SYGVVLWE 684

RESULT 74

T05335

hypothetical protein FIC12.190 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Sep-1999

C:Accession: T05335

R:Bevan, M.; Terry, N.; Ardiles, W.; Buysbaert, C.; Dasseville, R.; De Clerck, R.; De

ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15408

A:Accession: T05335

A:Molecule type: DNA

A:Residues: 1-992 <BEV>

A:Cross-references: EMBL:AL022224

A:Experimental source: cultivar Columbia; BAC clone FIC12

C:Genetics:

A:Map position: 4

A:Introns: 884/1

A:Note: FIC12.190

C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; F

Query Match 1.8%; Score 8; DB 2; Length 992;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVAVKKL 48

DB 734 EVAVKKL 741

RESULT 75

S49835

hypothetical protein YDR080w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D4446; hypothetical protein YD8554.13

C:Species: Saccharomyces cerevisiae

C:Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 29-Oct-1999

C:Accession: S49835; S48767; S55827; S67897

R:Richard, C.; Harris, D.E.

submitted to the EMBL Data Library, November 1994

A:Reference number: S49823

A:Accession: S49835

A:Molecule type: DNA

A:Residues: 1-992 <RIC>

A:Cross-references: EMBL:Z46796; NID:G577794; PIDN:CAA86802.1; PID:G577807

R:Coster, F.; Jonniaux, J.L.; Goffeau, A.

submitted to the EMBL Data Library, October 1994

A:Reference number: S48758

A:Accession: S48767

A:Molecule type: DNA

A:Residues: 1-423, 'W', 425-992 <COS>

A:Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251

R:Coster, F.; Jonniaux, J.L.; Goffeau, A.

Yeast 11, 673-679, 1995

A:Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr

A:Reference number: S55819; PMID:96093910; PMID:7483840

A:Accession: S55827

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-423, 'W', 425-992 <COM>

A:Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994

R:Foury, F.; Jonniaux, J.L.; Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67889

A:Accession: S67897

A:Molecule type: DNA

A:Residues: 1-992 <FOU>

A:Cross-references: EMBL:Z74376; NID:G1431548; PIDN:CAA98899.1; PID:G143154;

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:VPS41

A:Cross-references: SGD:S0002487; MIPS:YDR080w

A:Map position: 4R

Query Match 1.8%; Score 8; DB 2; Length 992;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443

DB 78 GDDDDDD 85

Search completed: December 5, 2003, 09:32:14

Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:21:47 ; Search time 17 seconds

(without alignments)
1258.656 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455
Sequence: 1 MSLSGSPYQIKPDLQFPE.....GDDDDDDGDEBDNDNDNSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	2.2	729	1 KEX1_YEAST	P09620 saccharomyc
2	10	2.2	1089	1 KEX1_YEAST	P38798 saccharomyc
3	9	2.0	375	1 MCK1_YEAST	P21965 saccharomyc
4	9	2.0	502	1 KIR3_MOUSE	O61288 mus musculu
5	9	2.0	505	1 KIR3_RAT	P80203 rattus norv
6	9	2.0	856	1 EZAI_ARATH	O92588 arabidopsi
7	9	2.0	1390	1 ILPR_BRALA	O02466 brachiosto
8	9	2.0	1390	1 INSR_AEDAE	O93105 aedes aegypt
9	8	1.8	157	1 UL51_HCMVA	P16792 human cytom
10	8	1.8	158	1 2SS_SOYEN	P19594 glycine max
11	8	1.8	229	1 REGO_BR82	P13870 bacterioph
12	8	1.8	286	1 MCM1_YEAST	P11746 saccharomyc
13	8	1.8	294	1 OSTP_MOUSE	P10923 mus musculu
14	8	1.8	317	1 OSTP_RAT	P08721 rattus norv
15	8	1.8	352	1 RLA0_HALCU	P17006 halobacteri
16	8	1.8	352	1 RLA0_HALCU	P13553 halobacteri
17	8	1.8	381	1 Y532_METUA	O57952 methanococ
18	8	1.8	471	1 SIDI_SCHPO	O14305 schizosacch
19	8	1.8	482	1 T2EA_YEAST	P36100 saccharomyc
20	8	1.8	487	1 STK4_HUMAN	O13043 homo sapien
21	8	1.8	491	1 STK4_HUMAN	O13188 homo sapien
22	8	1.8	579	1 M3K7_MOUSE	O62073 mus musculu
23	8	1.8	580	1 YW72_YEAST	O05021 saccharomyc
24	8	1.8	606	1 M3K7_HUMAN	O43318 homo sapien
25	8	1.8	746	1 E2H2_HUMAN	O15910 homo sapien
26	8	1.8	746	1 E2H2_MOUSE	O61188 mus musculu
27	8	1.8	799	1 CN3B_MOUSE	O61409 mus musculu
28	8	1.8	840	1 YH71_YEAST	P38835 saccharomyc
29	8	1.8	875	1 DD10_HUMAN	O13206 homo sapien
30	8	1.8	938	1 CAFA_HUMAN	O13111 homo sapien
31	8	1.8	943	1 ROR2_HUMAN	O01974 homo sapien
32	8	1.8	944	1 ROR2_MOUSE	O92138 mus musculu
33	8	1.8	992	1 VP41_YEAST	P38959 saccharomyc

34	8	1.8	1085	1 IPH1_YEAST	P39520 saccharomyc
35	8	1.8	1131	1 YN94_YEAST	P31380 saccharomyc
36	8	1.8	1165	1 YN94_YEAST	P53950 saccharomyc
37	7	1.5	86	1 RNPH_AZOVI	O91570 azotobacter
38	7	1.5	101	1 KOA2_ECOLI	P03052 escherichia
39	7	1.5	147	1 UCRH_YEAST	P00127 saccharomyc
40	7	1.5	160	1 HPPK_HAEIN	P43777 haemophilus
41	7	1.5	172	1 UBQ2_YEAST	P06104 saccharomyc
42	7	1.5	173	1 NUGM_BRABR	P43194 brachyramph
43	7	1.5	173	1 NUGM_BRABR	P43195 brachyramph
44	7	1.5	175	1 VIRR_AGRTU	O52278 agrobacteri
45	7	1.5	183	1 NO29_XENLA	O42584 xenopus lae
46	7	1.5	196	1 YE47_THEMA	O92149 thermocoga
47	7	1.5	204	1 HMG1_ONCMY	P07746 oncorhynch
48	7	1.5	207	1 HIT_MOUSE	O07133 mus musculu
49	7	1.5	207	1 RPOA_EUGAN	O85194 euglena ana
50	7	1.5	228	1 EF1B_ORYSA	O40680 oryza sativ
51	7	1.5	230	1 EF1B_BETUV	O81918 beta vulgar
52	7	1.5	245	1 ICP3_HSVIN	P37319 herpes simp
53	7	1.5	248	1 ICP3_HSV11	P36313 herpes simp
54	7	1.5	248	1 YFO4_SCHPO	O94611 schizosacch
55	7	1.5	250	1 YB95_ARATH	O8487 arabidopsi
56	7	1.5	252	1 ICP3_HSV1D	P37318 herpes simp
57	7	1.5	253	1 ASP_PLAFS	P13825 plasmodium
58	7	1.5	253	1 L133_CAEEL	P34684 caenorhabdi
59	7	1.5	253	1 T2D7_RAT	O62880 rattus norv
60	7	1.5	263	1 ICP3_HSV1P	P08353 herpes simp
61	7	1.5	264	1 OSTP_CHICK	P23498 gallus gall
62	7	1.5	264	1 T2D7_HUMAN	O16594 homo sapien
63	7	1.5	265	1 EF1D_XENLA	P29693 xenopus lae
64	7	1.5	268	1 A32E_HUMAN	O93870 homo sapien
65	7	1.5	280	1 HAY1_MOUSE	O35387 mus musculu
66	7	1.5	289	1 YN94_CAEEL	P34572 caenorhabdi
67	7	1.5	295	1 UBC3_YEAST	P14682 saccharomyc
68	7	1.5	296	1 PEX5_CANAL	O74711 candida alb
69	7	1.5	303	1 MURB_BACSU	P18579 bacillus su
70	7	1.5	305	1 CDK3_HUMAN	O00526 homo sapien
71	7	1.5	310	1 MLEP_YEAST	O92450 drosophila
72	7	1.5	314	1 MOK3_MOUSE	P54658 saccharomyc
73	7	1.5	314	1 MOK3_MOUSE	O63844 m mtogen-a
74	7	1.5	323	1 KRAF_MSV36	P00532 murine sarc
75	7	1.5	323	1 LUKE_STPAU	P31715 staphylococ
76	7	1.5	325	1 H1GB_STPAN	O07226 staphylococ
77	7	1.5	328	1 KRAE_MOUSE	P28028 mus musculu
78	7	1.5	332	1 AATP_WHEAT	O02066 triticum ae
79	7	1.5	332	1 CDC2_CAEEL	P34556 caenorhabdi
80	7	1.5	332	1 Y713_CHLAMU	O99173 chlamydia m
81	7	1.5	333	1 SPT2_YEAST	P06843 saccharomyc
82	7	1.5	341	1 S122_ARATH	O96922 arabidopsi
83	7	1.5	348	1 RLA0_HALMA	P15825 halarchula
84	7	1.5	349	1 STY1_SCHPO	O09692 schizosacch
85	7	1.5	353	1 ASK2_ARATH	P43292 arabidopsi
86	7	1.5	353	1 FUS3_YEAST	P16992 saccharomyc
87	7	1.5	355	1 ARG3_YEAST	P07250 saccharomyc
88	7	1.5	358	1 PKY1_HUMAN	P27703 mus musculu
89	7	1.5	358	1 PKY1_HUMAN	P51817 homo sapien
90	7	1.5	359	1 RMOS_XENLA	P12965 xenopus lae
91	7	1.5	360	1 MKO1_BOVIN	P46196 bos taurus
92	7	1.5	360	1 MKO1_HUMAN	P26842 homo sapien
93	7	1.5	361	1 MKO1_XENLA	P26656 xenopus lae
94	7	1.5	363	1 ASK1_ARATH	P43291 arabidopsi
95	7	1.5	365	1 M14E_DROME	O61443 drosophila
96	7	1.5	366	1 CPK1_CRYNE	O62618 cryptococcu
97	7	1.5	366	1 M14A_DROME	O62618 drosophila
98	7	1.5	367	1 MK12_MOUSE	O08911 mus musculu
99	7	1.5	367	1 MK12_MOUSE	O63538 rattus norv
100	7	1.5	367	1 RM1L_AVI11	P10533 avian retro

RESULT 1

ALIGNMENTS

KEX1_YEAST
 ID KEX1_YEAST STANDARD; PRT; 729 AA.
 AC P09620;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase KEX1 precursor (EC 3.4.16.6) (Carboxypeptidase D).
 GN KEX1 OR YGL303C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87273520; PubMed=3301004;
 RA Dmochowska A., Dignard D., Henning D., Thomas D.Y., Bussey H.;
 RT "Yeast KEX1 gene encodes a putative protease with a carboxypeptidase
 RT B-like function involved in killer toxin and alpha-factor precursor
 RT processing";
 RL Cell 50:573-584(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bjournson A.J., McReynolds A.D.K., Wright L.F.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=89343988; PubMed=2668738;
 RA Cooper A., Bussey H.;
 RT "Characterization of the yeast KEX1 gene product: a carboxypeptidase
 RT involved in processing secreted precursor proteins";
 RL Mol. Cell. Biol. 9:2706-2714(1989).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=96276234; PubMed=8745419;
 RA Shilton B.H., Li Y., Tessler D., Thomas D.Y., Cygler M.;
 RT "Crystallization of a soluble form of the Kex1 serine
 RT carboxypeptidase from Saccharomyces cerevisiae";
 RL Protein Sci. 5:395-397(1996).
 CC -!- FUNCTION: PROTEASE WITH A CARBOXYPEPTIDASE B-LIKE FUNCTION
 CC INVOLVED IN KILLER TOXIN AND ALPHA-FACTOR PRECURSOR PROCESSING.
 CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
 CC or lysine residue.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M17231; AAA34717.1; -.
 DR EMBL; 272725; CA96915.1; -.
 DR PIR; A29651; A29651.
 DR PDB; 1AC5; 15-MAY-97.
 DR MEROPS; S10.007; -.
 DR SGD; S0003171; KEX1.
 DR GO; GO:0005802; C:Golgi trans-face; IDA.
 DR InterPro; IPR000379; Ser esters site.
 DR InterPro; IPR001563; Serine catpept.
 DR Pfam; PF00450; serine catpept; 1.
 DR PRINTS; PR00724; CRBOXPTASEC.
 DR PRODOM; PD001189; serine catpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 KM Hydrolyase; Carboxypeptidase; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 729 CARBOXYPEPTIDASE KEX1.
 FT ACT_SITE 198 729
 FT ACT_SITE 405 405 BY SIMILARITY.
 FT ACT_SITE 470 470 BY SIMILARITY.

FT CARBOHYD 81
 FT CARBOHYD 459
 FT CARBOHYD 467
 FT DOMAIN 506
 FT TRANSMEM 611
 FT TRANSMEM 619
 FT TRANSMEM 637
 FT TRANSMEM 698
 FT HELIX 26
 FT HELIX 28
 FT TURN 29
 FT STRAND 30
 FT STRAND 33
 FT HELIX 35
 FT TURN 37
 FT HELIX 39
 FT TURN 41
 FT STRAND 45
 FT STRAND 50
 FT STRAND 57
 FT STRAND 72
 FT HELIX 84
 FT STRAND 86
 FT STRAND 91
 FT TURN 95
 FT STRAND 100
 FT STRAND 101
 FT TURN 103
 FT HELIX 104
 FT STRAND 109
 FT STRAND 114
 FT TURN 116
 FT STRAND 118
 FT STRAND 122
 FT TURN 126
 FT STRAND 127
 FT HELIX 129
 FT STRAND 132
 FT STRAND 134
 FT TURN 138
 FT TURN 142
 FT TURN 143
 FT TURN 145
 FT HELIX 154
 FT TURN 159
 FT HELIX 166
 FT TURN 185
 FT HELIX 187
 FT STRAND 191
 FT TURN 199
 FT STRAND 200
 FT HELIX 217
 FT TURN 221
 FT STRAND 226
 FT HELIX 239
 FT TURN 244
 FT HELIX 246
 FT TURN 253
 FT TURN 258
 FT TURN 261
 FT HELIX 263
 FT HELIX 283
 FT HELIX 291
 FT TURN 295
 FT HELIX 296
 FT TURN 304
 FT TURN 309
 FT TURN 310
 FT TURN 312
 FT STRAND 315
 FT TURN 318
 FT STRAND 320
 FT STRAND 327
 FT TURN 332
 FT TURN 336
 FT HELIX 346
 FT TURN 347
 FT TURN 348
 FT HELIX 349
 FT TURN 350
 FT TURN 354
 FT TURN 358
 FT STRAND 367
 FT STRAND 369
 FT HELIX 374
 FT HELIX 383
 FT HELIX 387
 FT TURN 393
 FT STRAND 396
 FT TURN 403
 FT TURN 404
 FT TURN 407
 FT TURN 408

N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .).
 N-LINKED (GLCNAC. . .).
 ASP/GLU-RICH (ACIDIC).
 POTENTIAL.
 S->A: INACTIVATES ENZYME.

```

FT HELIX 410 419
FT STRAND 421 422
RA 423 424
FT STRAND 425 426
FT STRAND 430 431
FT STRAND 433 439
FT STRAND 443 444
FT STRAND 452 457
FT STRAND 458 459
FT STRAND 460 465
FT STRAND 466 467
FT HELIX 472 475
FT HELIX 477 487
FT STRAND 488 489
FT STRAND 492 496
FT STRAND 497 498
FT STRAND 499 504
SQ SEQUENCE 729 AA; 82245 MW; 70583F279AC02A41 CRC64;

Query Match
Best Local Similarity 100.0%; Score 10; DB 1; Length 729;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 434 SEGDDDDDD 443
DB 564 SEGDDDDDD 573

RESULT 2
NMND2_YEAST
ID NMND2_YEAST STANDARD; PRT; 1089 AA.
AC P38798;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Nonsense-mediated mRNA decay protein 2 (Up-frameshift suppressor 2).
GN NMND2 OR UPF2 OR IRS1 OR SUW1 OR YHR077C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95189083; PubMed=7883168;
RT He F., Jacobson A.;
RT "Identification of a novel component of the nonsense-mediated mRNA
RT decay pathway by use of an interacting protein screen.";
RT Gene Dev. 9:437-454(1995).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=PLY136;
RC MEDLINE=95189082; PubMed=7883167;
RA Cui Y., Hagan K.W., Zhang S., Peliz S.W.;
RT "Identification and characterization of genes that are required for
RT the accelerated degradation of mRNAs containing a premature
RT translational termination codon.";
RT Gene Dev. 9:423-436(1995).
RL [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=95327692; PubMed=7604038;
RA Lee S.I., Umen J.G., Varnus H.E.;
RT "A genetic screen identifies cellular factors involved in retroviral
RT -1 frameshifting.";
RT Proc. Natl. Acad. Sci. U.S.A. 92:6587-6591(1995).
RL [4]
RN SEQUENCE FROM N.A.
RP MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latrille P., Louis E.J., Macri C., Mardis E., Meneses S., Mousier L.,
RA Nhan M., Ritkin L., Riles L., St Peter H., Trevaakis E., Vaughan K.,

```

```

RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII.";
RL Science 265:2077-2082(1994).
CC -1- FUNCTION: INVOLVED IN NONSENSE-MEDIATED DECAY OF MRNAS CONTAINING
CC PREMATURE STOP CODONS. IT INTERACTS, VIA ITS C-TERMINUS, WITH
CC NAM7/UPF1. COULD BE INVOLVED IN DETERMINING THE EFFICIENCY OF
CC TRANSLATIONAL TERMINATION OR REINITIATION OR FACTORS INVOLVED IN
CC THE INITIAL ASSEMBLY OF AN INITIATION- AND TERMINATION-COMPETENT
CC MRNP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U14974; AAA67724.1; -
DR EMBL; U12137; AAA66521.1; -
DR EMBL; U28158; AAA74948.1; -
DR EMBL; U10556; AAA68893.1; -
DR PIR; S48244; S48244.
DR SGD; S000119; NMND2.
DR GO; GO:0005844; C:Polysome; IDA.
DR InterPro; IPR003890; IF_eIF4G.
DR Pfam; PF02854; MIF4G; 3.
DR Pfam; PF04050; Upf2; 1.
DR SMART; SM00543; MIF4G; 3.
FT DOMAIN 843 975 ASP/GLU-RICH (HIGHLY ACIDIC).
FT CONFILICT 2 D -> YQQ (IN REF. 3 AND 4).
SQ SEQUENCE 1089 AA; 126746 MW; 13BBE725675CBF52 CRC64;

Query Match
Best Local Similarity 100.0%; Score 10; DB 1; Length 1089;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446
DB 902 DDDDDDDGEE 911

RESULT 3
MCK1_YEAST
ID MCK1_YEAST STANDARD; PRT; 375 AA.
AC P21965;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein kinase MCK1 (EC 2.7.1.-) (Meiosis and centromere regulatory
DE kinase).
GN MCK1 OR YPK1 OR YNL307C OR N0392.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RP MEDLINE=91061730; PubMed=1701015;
RA Dailey D., Schieven G.L., Lim M.-Y., Marguardt H., Gilmore T.,
RA Thorner J., Martin G.S.;
RT "Novel yeast protein kinase (YPK1 gene product) is a 40-kilodalton
RT phosphotyrosyl protein kinase associated with protein-tyrosine kinase
RT activity.";
RT Mol. Cell. Biol. 10:6244-6256(1990).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=9184610; PubMed=2010083;
RA Neigeborn L., Mitchell A.P.;
RT "The yeast MCK1 gene encodes a protein kinase homolog that activates

```

RT early meiotic gene expression." ;
 RL Genes Dev. 5:533-548(1991).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91184611; PubMed=2010084;
 RA Shero J.H., Hietter P.;
 RT "A suppressor of a centromere DNA mutation encodes a putative protein
 kinase (MCK1)." ;
 RL Genes Dev. 5:549-560(1991).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1676;
 RX MEDLINE=96076632; PubMed=7502583;
 RA Maftchi M., Nicoud J.-M., Levesque H., Galliard C.;
 RT "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV
 identifies six known genes, a new member of the hexose transporter
 family and ten new open reading frames." ;
 RL Yeast 11:1077-1085(1995).
 RN (5)
 RP SEQUENCE OF 1-16 FROM N.A.
 RC STRAIN=S288c / FY1679;
 RX MEDLINE=96132033; PubMed=8553702;
 RA Maurer K.C.T., Urbanus J.H.M., Planta R.J.;
 RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV
 carrying a ribosomal protein gene cluster: the genes encoding a
 plasma membrane protein and a subunit of replication factor C, and a
 novel putative serine/threonine protein kinase gene." ;
 RL Yeast 11:1303-1310(1995).
 RN (1)
 RP FUNCTION: MAY BE AN AUTOPHOSPHORYLATING TYROSINE KINASE, A
 BIFUNCTIONAL (SERINE/TYROSINE-SPECIFIC) PROTEIN KINASE, OR A
 SERINE KINASE THAT IS A SUBSTRATE FOR AN ASSOCIATED TYROSINE
 KINASE. MCK1 IS A TRANSCRIPTIONAL ACTIVATOR OF IME1, IT STIMULATES
 SPORE MATURATION, AND PLAY A POSITIVE REGULATORY ROLE IN BOTH
 MITOTIC CENTROMERE FUNCTION AND ACTIVATION OF EARLY MEIOTIC GENE
 EXPRESSION.
 CC -1- PTM: PHOSPHORYLATED AT TYROSINE AND SERINE IN VIVO.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X55054; CAA38895.1; -
 CC EMBL: M55984; AAA34764.1; -
 CC EMBL: Z46259; CAA63388.1; -
 CC EMBL: Z71583; CAA96236.1; -
 CC EMBL: Z71582; CAA96235.1; -
 CC PIR: A39622; A39622.
 CC HSSP: P24941; JA01.
 CC SGD: S0005251; MCK1.
 CC GO: GO:0005625; F:protein threonine/tyrosine kinase activity; IDA.
 CC GO: GO:0004712; F:protein threonine/tyrosine kinase activity; IDA.
 CC GO: GO:0007126; F:mitosis; IMP.
 CC GO: GO:0000070; P:mitotic chromosome segregation; IGI.
 CC GO: GO:0006168; P:protein amino acid phosphorylation; IDA.
 CC GO: GO:0006950; P:response to stress; IGI.
 CC GO: GO:0007151; P:sporulation (sensu Saccharomycetes); IMP.
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR002290; Ser_thr_kinase.
 CC Pfam: PF00069; Pkinase; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SMO0220; S_TKC; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 35 327 PROTEIN KINASE.
 FT NP_BIND 41 49 ATP (BY SIMILARITY).

FT BINDING 68 68 ATP (BY SIMILARITY).
 FT ACT_SITE 164 164 BY SIMILARITY.
 FT MOD_RES 199 199 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 375 AA; 43136 MW; BCF0C776B8E3841B CRC64;
 Query Match 2.0%; Score 9; DB 1; Length 375;
 Best local Similarity 100.0%; Pred. No. 0.65;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 145 GVLTICDFG 153
 DB 177 GVLTICDFG 185
 RESULT 4
 KIR3_MOUSE STANDARD; PRT; 502 AA.
 ID KIR3_MOUSE
 AC Q61288; Q61289;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase receptor R3 precursor (EC 2.7.1.37)
 DE (SKR3) (Activin receptor-like kinase 1) (ALK-1) (TGF-B superfamily
 DE receptor type I) (TSR-I).
 GN ACVRL1 OR ACVRLK1 OR ALK-1.
 OS Mus musculus (Mouse).
 OC Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=96067535; PubMed=7488127;
 RA Wu X., Robinson C.E., Fong H.W., Crabtree J.S., Rodriguez B.R.,
 RA Roe B.A., Gimble J.M.;
 RT "Cloning and characterization of the murine activin receptor like
 RT kinase-1 (ALK-1) homolog." ;
 RL Biochem. Biophys. Res. Commun. 216:78-83(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=95269711; PubMed=7750489;
 RA Delfin K., Verschueren K., Lonnoy O., Moren A., Grimsby S.,
 RA Spiegler K., Myazano K., Huybreck D., ten Dijke P.;
 RT "Distinct spatial and temporal expression patterns of two type I
 RT receptors for bone morphogenetic proteins during mouse
 RT embryogenesis." ;
 RL Endocrinology 136:2652-2663(1995).
 CC -1- FUNCTION: TYPE I/TYPE II TGF-BETA RECEPTORS FORM AN HETEROMERIC
 CC COMPLEX AFTER BINDING TGF-BETA AT THE CELL SURFACE AND ACT AS
 CC SIGNAL TRANSDUCERS. MAY BIND ACTIVIN AS WELL.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC TGFb RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L48015; AAB03642.1; -
 CC EMBL: Z31664; CAA83484.1; -
 CC PIR: I48241; I48241.
 CC PIR: JC4337; JC4337.
 CC WGI: I338946; ACVRL1.
 CC InterPro: IPR000472; Activin_rec.
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR002290; Ser_thr_kinase.
 CC InterPro: IPR003605; TGFbeta_GS.

DR Pfam; PF01064; Activin_recpt; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 502
 FT
 FT DOMAIN 23 119
 FT TRANSMEM 120 140
 FT DOMAIN 141 502
 FT DOMAIN 201 502
 FT NP_BIND 208 215
 FT BINDING 228 228
 FT ACT_SITE 329 329
 FT CARBOHYD 32 32
 FT CARBOHYD 97 97
 FT CONFLICT 17 17
 FT CONFLICT 21 21
 FT CONFLICT 23 23
 FT CONFLICT 25 25
 FT CONFLICT 305 305
 FT CONFLICT 358 358
 FT CONFLICT 366 366
 SQ SEQUENCE 502 AA; 56632 MW; 092AB29778BB004 CRC64;

Query Match
 Best Local Similarity 2.0%; Score 9; DB 1; Length 502;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139
 Db 327 HRDLKSRNV 335

RESULT 5
 KIR3_RAT STANDARD; PRT; 505 AA.
 AC P80203;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase receptor R3 precursor (EC 2.7.1.37)
 GN (SKR3) (TGF-B superfamily receptor type I) (TSR-I).
 GN ACVR1 OR ACVR1L
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Urogenital ridge;
 RX MEDLINE=93372378; PubMed=8395916;
 RA He W.-W., Gustafson M., Hirobe S., Donahoe P.;
 RT "Developmental expression of four novel serine/threonine kinase
 RT receptors homologous to the activin/transforming growth factor-beta
 RT type II receptor family";
 RL Dev. Dyn. 196;133-142(1993).
 CC -1- FUNCTION: TYPE I/TYPE II TGF-BETA RECEPTORS FORM AN HETEROMERIC
 CC COMPLEX AFTER BINDING TGF-BETA AT THE CELL SURFACE AND ACT AS
 CC SIGNAL TRANSDUCERS. MAY BIND ACTIVIN AS WELL.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: UROGENITAL RIDGE, TESTIS, OVARY, BRAIN AND
 CC LUNGS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC TGFβ RECEPTOR SUBFAMILY.
 DR InterPro; IPR000472; Activin_rec.
 DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR003605; TGFbeta_GS.
 DR Pfam; PF01064; Activin_recpt; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 505
 FT
 FT DOMAIN 21 121
 FT TRANSMEM 122 142
 FT DOMAIN 143 505
 FT DOMAIN 203 505
 FT NP_BIND 209 217
 FT BINDING 230 230
 FT ACT_SITE 332 332
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 SQ SEQUENCE 505 AA; 56774 MW; 4C10D09C9105F23C CRC64;

Query Match
 Best Local Similarity 2.0%; Score 9; DB 1; Length 505;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139
 Db 330 HRDLKSRNV 338

RESULT 6
 EZAL_ARATH STANDARD; PRT; 856 AA.
 AC Q9ZSH8; O04246;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Potential Polycomb group protein EZAL (CURLY LEAF-like 1).
 GN EZAL OR ATAG02020 OR T10W13.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bilodeau P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;
 RT "EZAL, a novel polycomb group gene from Arabidopsis thaliana";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Scheller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoef A., Stiekema W., Ertlan K.-D., Terry N.,
 RA Harris B., Anorg W., Brandt P., Griwell L.A., Rieger M.,
 RA Weichselgartner W., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Deleney M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.A., McCullagh B., Bilham L., Robben J.,
 RA van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbusche F.,
 RA Braeken M., Weltjens I., Vost M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzengger T., Botne G., Ransperger U., Hilbert H., Braun M.,
 RA Holzner B., Brandt A., Peters S., van Staveren M., Dirse W.,
 RA Woolfman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,

Clark L., Doggett J., Hall S., Kay M., Lennard N., McClay K., Mayes R.,
 Pettit A., Rajadream M.A., Lyne M., Bens V., Rechmann S.,
 Borckova D., Blocher H., Scharle M., Grimm M., Loehner T.H.,
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fartmann B., Grandenach K., Danner D., Herzl A.,
 Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 Messner O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 Cheifor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 Gibbons T., Weber N., Vandembol M., Barges M., Teoi J., Torres A.,
 Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacou D., Jesse T.,
 Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bietke C.,
 Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 Zaccaria P., Bevan M., Wilson R.K., de la Baetide M., Habermann K.,
 Barnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 Sekhon M., Murray J., Sheet P., Cordes M., Abu-Itreiden J.,
 Stoneking T., Kallunki J., Graves T., Harmon G., Edwards J.,
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 Nelson J., Spieth J., Ryan E., Andrews S., Gelsel C., Layman D.,
 Du H., Ali J., Bergtholt A., Jones K., Drone K., Cotton M., Joshua C.,
 Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matere A., Shah R.,
 Shaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 Grant S., Shohdy N., Hasegawa A., Hamed A., Lohi M., Johnson A.,
 Chen B., Marra M., Martensen R., McCombe W.R.,
 "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 Nature 402:769-777(1999).
 [3]
 SEQUENCE FROM N.A.
 STRAIN=cy. Columbia;
 Srinivasak K., Davis R.W., Ecker J.R., Theologis A.,
 "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 SFP consortium (Salk/Stanford/PGEC).";
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: Potential polycomb group (PcG) protein. PcG proteins act
 by forming multiprotein complexes, which are required to maintain
 the transcriptionally repressive state of homeotic genes
 throughout development. PcG proteins are not required to initiate
 repression, but to maintain it during later stages of development.
 They probably act via the methylation of histones, rendering
 chromatin heritably changed in its expressibility. Its sequence
 suggests that it may participate in a potential methyltransferase
 activity of PcG complexes (by similarity).
 -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 -1- SIMILARITY: BELONGS TO THE EZ FAMILY.
 -1- SIMILARITY: CONTAINS 1 SANT DOMAIN.
 -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
 -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 gene model prediction.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: AF100163; AAD09108.1; -
 EMBL: AF001308; AAC78694.1; ALT_SEQ.
 EMBL: AL161493; CAB80695.1; ALT_SEQ.
 EMBL: AY057477; AAL09711.1; -
 EMBL: AY090293; AAL09054.1; -
 PIR: T52415; T52415.
 HSSP: P10969; 1WGT.
 InterPro: IPR001005; Myb_DNA_binding.
 InterPro: IPR001214; SET.
 Pfam: PF00856; SET; 1.
 SMART: SM00717; SANT; 1.
 SMART: SM00317; SET; 1.

DR PROSITE; PSS0280; SET; 1.
 KW Transcription regulation; Repressor; Nuclear protein.
 FT DOMAIN 489 539 SANT.
 FT DOMAIN 706 826 SET.
 FT DOMAIN 605 680 CYS-RICH.
 FT DOMAIN 17 33 ASP/GLU-RICH (ACIDIC).
 SO SEQUENCE 856 AA; 95396 MW; DDA809C936F197C CRC64;
 Query Match 2.0%; Score 9; DB 1; Length 856;
 Best Local Similarity 100.0%; Fred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Gaps 0;
 Oy 439 DDDDDGEE 447
 Db 17 DDDDDGEE 25
 RESULT 7
 ILPR_BRALA STANDARD; PRT; 1363 AA.
 ID ILPR_BRALA
 AC 002466;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).
 OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
 OC Branchiostoma.
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 NCBI_TaxID=7740;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96408719; PubMed=8813726;
 RA Pashmforough M., Chan S.-J., Steiner D.F.,
 RT "Structure and expression of the insulin-like peptide receptor from
 amphioxus.";
 RL Mol. Endocrinol. 10:857-866(1996).
 CC -1- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND
 HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -1- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY
 DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF
 THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE
 DOMAIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 RECEPTOR SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: S83394; AAB50848.1; -
 PIR: T43220; T43220.
 HSSP: P06213; 1IRK.
 InterPro: IPR000494; EGFR_L domain.
 InterPro: IPR003961; FN_III-
 InterPro: IPR006211; Furin-like.
 InterPro: IPR006212; Furin-repeat.
 InterPro: IPR000719; Prot_Kinase.
 InterPro: IPR002011; RTKinaseII.
 InterPro: IPR001245; Tyr_kinase.
 Pfam: PF00041; fn3; 3.
 Pfam: PF00757; Furin-like; 1.
 Pfam: PF00069; kinase; 1.
 Pfam: PF01030; Recep_L domain; 2.
 PRINTS: PR0109; TYRKINASE.
 ProDom: PD000001; Prot_kinase; 1.
 SMART: SM00060; FN3; 3.

DR SMART; SM00261; FU; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00239; RECEPTOR TYR_KIN_II; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KM Glycoprotein; ATP-binding; Phosphorylation; Signal.
 FT SGNL 1 29
 FT CHAIN 30 716
 FT PROPEP 717 720
 FT CHAIN 721 1363
 FT DOMAIN 950 1363
 FT NP_BIND 994 1283
 FT BINDING 1000 1008
 FT ACT_SITE 1148 1148
 FT MOD_RES 1174 1174
 FT CARBOHYD 51 51
 FT CARBOHYD 97 97
 FT CARBOHYD 137 137
 FT CARBOHYD 278 278
 FT CARBOHYD 483 483
 FT CARBOHYD 599 599
 FT CARBOHYD 617 617
 FT CARBOHYD 665 665
 FT CARBOHYD 711 711
 FT CARBOHYD 732 732
 FT CARBOHYD 736 736
 FT CARBOHYD 743 743
 FT CARBOHYD 816 816
 FT CARBOHYD 885 885
 FT CARBOHYD 898 898
 SQ SEQUENCE 1363 AA; 154104 MW; 228120B4EAB1ED65 CRC64;

Query Match 2.0%; Score 9; DB 1; Length 1363;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVIVMEM 199
 DB 1210 SYGVIVMEM 1218

RESULT 8
 INSR_AEDAE STANDARD; PRT; 1390 AA.
 ID INSR_AEDAE
 AC Q93105;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin-like receptor precursor (EC 2.7.1.112) (MIR).
 OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UGAL; TISSUE=Ovary;
 RA MEDLINE=97254344; PubMed=9099579;
 RA Graf R., Neuenschwander S., Brown M.R., Ackermann U.;
 RT "Insulin-mediated secretion of ecdysteroids from mosquito ovaries and
 RT molecular cloning of the insulin receptor homologue from ovaries of
 RT blooded Aedes aegypti";
 RL Insect Mol. Biol. 6:151-163 (1997).
 CC -1- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN
 CC AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphate.
 CC -1- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
 CC bonds. The alpha chains contribute to the formation of the ligand-
 CC binding domain, while the beta chains carry the kinase domain (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U72939; AAB17094.1; -.
 CC PIR; T30346; T30346.
 DR HSSP; P06213; 1TRK.
 DR InterPro; IPR000494; EGFR_L_domain.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin repeat.
 DR InterPro; IPR000719; Prot Kinase.
 DR InterPro; IPR002011; RTKaseit.
 DR InterPro; IPR001245; Tyr_PKinase.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR PRINTS; PRO0109; TYRKINASE.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00261; FU; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00239; RECEPTOR TYR_KIN_II; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KM Glycoprotein; ATP-binding; Phosphorylation; Signal.
 FT SGNL 1 19
 FT CHAIN 20 753
 FT PROPEP 754 757
 FT CHAIN 758 1390
 FT DOMAIN 967 966
 FT TRANSMEM 967 987
 FT DOMAIN 988 1390
 FT NP_BIND 1025 1301
 FT BINDING 1031 1039
 FT ACT_SITE 1059 1059
 FT MOD_RES 1162 1162
 FT CARBOHYD 76 76
 FT CARBOHYD 196 196
 FT CARBOHYD 257 257
 FT CARBOHYD 332 332
 FT CARBOHYD 340 340
 FT CARBOHYD 439 439
 FT CARBOHYD 457 457
 FT CARBOHYD 492 492
 FT CARBOHYD 553 553
 FT CARBOHYD 765 765
 FT CARBOHYD 772 772
 FT CARBOHYD 793 793
 FT CARBOHYD 796 796
 FT CARBOHYD 871 871
 FT CARBOHYD 919 919
 FT CARBOHYD 933 933
 SQ SEQUENCE 1390 AA; 156631 MW; 6BD2AA382EFD6442 CRC64;

Query Match 2.0%; Score 9; DB 1; Length 1390;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVLMEM 199
 |||||
 Db 1224 SYGVLMEM 1232

RESULT 9
 UL51_HCMVA STANDARD; PRT; 157 AA.
 AC P16752;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Hypochemical protein UL51.
 GN UL51.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Hornslell T., Hutchison C.A. III, Kouzides T., Martignetti J.A.,
 RA Pridmore E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169. "
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X17403; CAJ35410.1; -.
 DR PIR: S09814; S09814.
 DR InterPro: IPR005208; Herpes_UL33.
 DR Pfam: PF03581; Herpes_UL33; 1.
 KW Hypochemical protein.
 SQ SEQUENCE 157 AA; 16968 MW; 5999035AC654517D CRC64;

Query Match 1.8%; Score 8; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 DDDGGEE 447
 |||||
 Db 12 DDDGGEE 19

RESULT 10
 2SS_SOYBN STANDARD; PRT; 158 AA.
 AC P19594;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 2S albumin precursor (GMS-1) [Contains: Aspartic acid-rich peptide;
 DE 8 kDa methionine-rich protein (8 kDa MRP)].
 GN Glycine max (Soybean).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=CV, Hodgson 78; TISSUE=Cotyledon;
 RA Galvez A.F., Revilla M.J.R., de Lumen B.O.;
 RT "A novel methionine-rich protein from soybean cotyledon: cloning and
 RT characterization of cDNA."
 RL (in) Plant Gene Register PGR97-103.
 RN [12]
 RP SEQUENCE OF 22-64.

RC TISSUE=Seed.
 RX MEDLINE=87280104; PubMed=3611081;
 RA Odani S., Koide T., Ono T.;
 RT "Amino acid sequence of a soybean (Glycine max) seed polypeptide
 RT having a poly(L-aspartic acid) structure."
 RL J. Biol. Chem. 262:10502-10505(1987).

RN [3]
 RP SEQUENCE OF 82-96.
 RA Revilla M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;
 RT "An 8 kDa methionine-rich protein (MRP) from soybean (Glycine max)
 RT cotyledon: identification, purification and N-terminal sequence."
 RL J. Agric. Food Chem. 44:2930-2935(1996).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE PROTEIN CONSISTS OF TWO CHAINS LINKED BY 2 DISULFIDE
 CC BONDS.

CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF005030; AAB71140.1; -.
 DR PIR: T05710; T05710.
 DR InterPro: IPR003612; AAI.
 DR Pfam: PF00234; tryp_alpha_aml; 1.
 DR SMART: SM00499; AAI; 1.
 KW Signal; Seed storage protein.
 KM SIGNAL; Seed storage protein.
 FT SIGNAL 1 21
 FT CHAIN 22 64

FT 2S ALBUMIN, SMALL CHAIN (ASPARTIC ACID-
 FT RICH PEPTIDE).
 FT PROPEP 65 81
 FT CHAIN 82 158
 FT SITE 54 56
 FT DOMAIN 56 64
 FT DOMAIN 88 91
 FT POLY-ASP.
 FT POLY-GLU.
 SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 158;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 |||||
 Db 55 GDDDDDD 62

RESULT 11
 REGQ_BP82 STANDARD; PRT; 229 AA.
 ID REGQ_BP82
 AC P13870;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antitermination protein Q.
 GN Q.
 OS Bacteriophage 82.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 OX NCBI_TaxID=10705;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87308148; PubMed=3624233;

RA Goliger J.A., Roberts J.W.;
 RT "Bacteriophage 82 gene Q and Q protein. Sequence, overproduction, and
 RT activity as a transcription antiterminator in vitro.";
 RL J. Biol. Chem. 262:11721-11725(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96196428; PubMed=8648624;
 RA Mahdi A.A., Sharples G.T., Mandal T.N., Lloyd R.G.;
 RT "Holiday junction resolvases encoded by homologous rna genes in
 RT *Escherichia coli* K-12 and phage 82.";
 RL J. Mol. Biol. 257:561-573(1996).
 CC -1- FUNCTION: POSITIVELY REGULATE EXPRESSION OF THE PHAGE LATE GENE
 CC OPERONS. BACTERIAL HOST RNA POLYMERASE MODIFIED BY ANTITERMINATION
 CC PROTEINS TRANSCRIBES THROUGH TERMINATION SITES THAT OTHERWISE
 CC PREVENT EXPRESSION OF THE REGULATED GENES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02803; AAA32298.1; -
 DR EMBL; X92588; CAA63332.1; -
 DR PIR; A29791; PGBP82.
 KW Transcription regulation; Transcription termination; DNA-binding.
 FT DNA BIND 14 32 POTENTIAL.
 FT SEQUENCE 229 AA; 26400 MW; E6BDABE818498667 CRC64;
 SQ
 Query Match 1.8%; Score 8; DB 1; Length 229;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKKL 297
 Db 142 TLERLKKL 149

RESULT 12
 MCM1 YEAST
 ID MCM1 YEAST STANDARD; PRT; 286 AA.
 AC P11746;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phenomone receptor transcription factor (SRM/PRTF protein).
 GN MCM1 OR FUN80 OR YMR043W OR YME532.08.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 OC NCB1 TaxID=4932;
 OX (1)
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=90249735; PubMed=2159934;
 RA Ammerer G.;
 RT "Identification, purification, and cloning of a polypeptide
 RT (PRTF/GM) that binds to mating-specific promoter elements in
 RT yeast.";
 RL Gene Dev. 4:299-312(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89141759; PubMed=3066908;
 RA Passmore S., Maine G.T., Elbie R., Christ C., Tye B.K.;
 RT "Saccharomyces cerevisiae protein involved in plasmid maintenance is
 RT necessary for mating of MAT alpha cells";
 RL J. Mol. Biol. 204:593-606(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8030692; PubMed=3311883;
 RA Dubois E., Bercy J., Descamps F., Messenguy F.;
 RT "Characterization of two new genes essential for vegetative growth in

RT *Saccharomyces cerevisiae*: nucleotide sequence determination and
 RT chromosome mapping";
 RL Gene 55:265-275(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Taggart K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
 RT XIII.";
 RL Nature 387:90-93(1997).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF COMPLEX WITH MAT2.
 RX MEDLINE=98140702; PubMed=9490409;
 RA Tan S., Richmond T.J.;
 RT "Crystal structure of the yeast Mat2alpha/MCM1/DNA ternary complex.";
 RL Nature 391:660-666(1998).
 CC -1- FUNCTION: INTERACTS WITH THE ALPHA-2 REPRESSOR OR WITH THE
 CC ALPHA-1 ACTIVATOR THEREBY REGULATING THE EXPRESSION OF MATING-
 CC TYPE-SPECIFIC GENES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
 CC FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52453; CAA36691.1; -
 DR EMBL; X14187; CAA32389.1; -
 DR EMBL; M17511; AAA34609.1; -
 DR EMBL; Z48502; CAA88409.1; -
 DR PIR; A34599; A34599.
 DR PDB; 1NMW; 18-MAR-98.
 DR TRANSFAC; T00500; -
 DR TRANSFAC; T00501; -
 DR SGD; S0004645; MCM1.
 DR GO; GO:0005634; C:nucleus; IPT.
 DR InterPro; IPR002100; TF_MADSbox.
 DR Pfam; PF00319; SRP-TE; 1.
 DR PRINTS; PR00404; MADSDOMAIN.
 DR SMART; SM00432; MADS; 1.
 DR PROSITE; PS00350; MADS_BOX_1; 1.
 DR PROSITE; PS50066; MADS_BOX_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.
 FT DOMAIN 18 72
 FT COMPACT 98 120
 FT COMPACT 9 9
 FT COMPACT 37 37
 FT COMPACT 156 157
 FT COMPACT 158 286
 FT HELIX 29 54
 FT TURN 55 55
 FT STRAND 58 63
 FT TURN 65 66
 FT STRAND 69 73
 FT HELIX 75 77
 FT STRAND 78 82
 FT HELIX 84 93
 FT TURN 94 94
 SQ SEQUENCE 286 AA; 32802 MW; FD75B4BF549B9B3B CRC64;
 Query Match 1.8%; Score 8; DB 1; Length 286;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 1) (SPP-1).
GN SPP1 OR SPP-1 OR 2B7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93054745; PubMed=1429723;
RA Singh K., Mukherjee A.B., de Vonge M.W., Mukherjee B.B.;
RT "Differential processing of osteopontin transcripts in rat kidney-
RT and osteoblast-derived cell lines";
RL J. Biol. Chem. 267:23847-23851(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067405; PubMed=3024151;
RA Oldberg A., Franzen A., Heinegaard D.;
RT "Cloning and sequence analysis of rat bone sialoprotein (osteopontin)
RT cDNA reveals an Arg-Gly-Asp cell-binding sequence";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8819-8823(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Smooth muscle;
RC MEDLINE=94013467; PubMed=8408622;
RA Giachelli C.M., Bae N., Almeida M., Denhardt D.T., Alpers C.E.,
RA Schwartz S.M.;
RT "Osteopontin is elevated during neointima formation in rat arteries
RT and is a novel component of human atherosclerotic plaques";
RL J. Clin. Invest. 92:1686-1696(1993).
CC -1- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an
CC integral part of the mineralized matrix. Probably important to
CC cell-matrix interaction.
CC -1- FUNCTION: Acts as a cytokine involved in enhancing production of
CC interferon-gamma and interleukin-12 and reducing production of
CC interleukin-10 and is essential in the pathway that leads to type
CC I immunity (By similarity).
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.
CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M99252; AAAA1765.1; -;
DR EMBL: M14656; AAAA1762.1; -;
DR PIR: A25917; A25917.
DR GO: GO:0005578; C:extracellular matrix; ISS.
DR GO: GO:0008189; F:apoptosis inhibitor activity; ISS.
DR GO: GO:0042056; F:chemoattractant activity; ISS.
DR GO: GO:0005125; F:cytokine activity; ISS.
DR GO: GO:0003793; F:defense/immunity protein activity; ISS.
DR GO: GO:0008083; F:growth factor activity; ISS.
DR GO: GO:0005178; F:integrin binding activity; ISS.
DR GO: GO:0006916; P:anti-apoptosis; ISS.
DR GO: GO:0030282; P:bone mineralization; ISS.
DR GO: GO:0007267; P:cell-cell signaling; ISS.
DR GO: GO:0007160; P:cell-matrix adhesion; ISS.
DR GO: GO:0030595; P:immune cell chemotaxis; ISS.
DR GO: GO:0006954; P:inflammatory response; ISS.
DR GO: GO:0042102; P:positive regulation of T-cell proliferation; ISS.
DR GO: GO:0045637; P:regulation of myeloid blood cell differentiation; ISS.
DR GO: GO:0042088; P:t-helper 1 type immune response; ISS.
DR Pfam: PF00865; Osteopontin; 1.
DR PRINTS; PR00216; OSTEOPONTIN.

DR SMART; SM00017; OSTEO; 1.
DR PROSITE; PS00884; OSTEOPONTIN; 1.
KW Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;
KW Phosphorylation; Signal.
FT SIGNAL 1 22
FT CHAIN 23 317
FT DOMAIN 86 96
FT SITE 144 146
FT CARBOHYD 79 79
FT CONFLICT 8
SQ SEQUENCE 317 AA; 34963 MW; 73CB5C21FF62310 CRC64;
QY 437 DDDDDDDG 444
Db 88 DDDDDDDG 95
Query Match 1.8%; Score 8; DB 1; Length 317;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 15
RLA0 HALCU STANDARD; PRT; 352 AA.
ID RLA0 HALCU
AC P17006;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acidic ribosomal protein P0 homolog (L10E).
GN RPLP0.
OS Halobacterium cutribnum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=34001;
RX MEDLINE=89305527; PubMed=2743981;
RA Shitman L.C., Dennis P.P.;
RT "Characterization of the L11, L1, L10 and L12 equivalent ribosomal
RT protein gene cluster of the halophilic archaeobacterium Halobacterium
RT cutribnum";
RL EMBO J. 8:1225-1235(1989).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15078; CAA33180.1; -;
DR HAMAP; MF_00280; -; 1.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Ribosomal protein.
SQ SEQUENCE 352 AA; 37199 MW; 2AB62E79A98D02D CRC64;
QY 437 DDDDDDDG 444
Db 333 DDDDDDDG 340
Query Match 1.8%; Score 8; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 16
RLA0 HALN1

```

ID RL00 HALNI STANDARD; PRT; 352 AA.
AC P13553; O9H0L4:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acidic ribosomal protein P0 homolog (L10E).
GN RPLP0 OR RPL10P OR VNG1104G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OC Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
NCBI_TaxID=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Halobium; STRAIN=9;
RX MEDLINE=88329082; PubMed=2458258;
RA Itoh T.;
RT "Complete nucleotide sequence of the ribosomal 'A' protein operon
RL from the archaeobacterium, Halobacterium halobium.";
RL Eur. J. Biochem. 176:297-303(1988).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lesky S.R., Baliga N.S., Thorsen V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weitz R., Goo Y.A.,
RA Leitauer B., Keller K., Cruz R., Danson M.U., Hough D.W.,
RA Maddocks D.G., Jablonki P.E., Krebs M.P., Angewine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschoder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dasarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RC Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13008; CAA31431.1; -.
DR EMBL; AE005042; AAC19499.1; -.
DR PIR; G84266; G84266.
DR PIR; S01315; RS5SL0.
DR HAMAP; MF_00280; -.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 352 AA; 37213 MW; E2F088CF34FDB9BD CRC64;

```

```

Query Match 1.8%; Score 8; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 17
ID Y532 METUA STANDARD; PRT; 391 AA.
AC 057952;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0532 precursor.

```

```

GN M0532.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Flicger L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fushman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurest M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE CHLP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67502; AAB98523.1; -.
DR PIR; D64366; D64366.
DR TIGR; M0532;
DR InterPro; IPR001327; PAD_Pyr_redox.
DR InterPro; IPR000733; Flav_monooxygenase.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR003042; Rng_monooxygenase.
DR Pfam; PF01360; Monooxygenase; 1.
DR PRINTS; PR00368; PADPXR.
DR PRINTS; PR00411; FNDRTASEI.
DR PRINTS; PR00420; RINGMONOXGNASE.
KW Hypothetical protein; signal; Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 391
SQ SEQUENCE 391 AA; 43094 MW; A2F69AEBF940BF2 CRC64;

```

```

Query Match 1.8%; Score 8; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 18
ID S1D1 SCHPO STANDARD; PRT; 471 AA.
AC 014305;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase s1d1 (EC 2.7.1.37) (STE20-like kinase
DE s1d1).
GN S1D1 OR SPAC9G1.09.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=21559449; PubMed=11384993;

```

RA Guerlin D.A., McColllum D.;
 RT "interaction between the noncatalytic region of Sidp kinase and
 RT Cdc14p is required for full catalytic activity and localization of
 RT Sidp.";
 RL J. Biol. Chem. 276:28185-28189 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21648401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moutie S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussey D., Barrell B.G., Nurse P.;
 RT "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880 (2002).
 RN [3]
 RN FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RP MEDLINE=20237617; PubMed=10775265; Gould K.L., McColllum D.;
 RX Guerlin D.A., Chang L., Irish F., Gould K.L., McColllum D.;
 RT "The role of the sidp kinase and cdc14p in regulating the onset of
 RT cytokinesis in fission yeast.";
 RL EMO J. 19:1803-1815 (2000).
 CC -1- FUNCTION: Has a role in the septation initiation network (SIN)
 CC -1- REQUIRED FOR CYTOKINESIS.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBUNIT: Interacts with cdc14.
 CC -1- SUBCELLULAR LOCATION: Spindle pole body. Localizes to the SPB
 CC prior to cytokinesis and leaves once septation is complete.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC STEREO SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 298763; CAB11493.1; -
 DR PIR: T19232; T19232.
 DR HSSP: 000534; 1B18.
 DR GeneDB: SPombe:SPAC9G1.09; -
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

KW Cell cycle; Cell division; Mitosis; Transferase; Kinase; ATP-binding.
 SQ SEQUENCE 471 AA; 52966 MW; 7332F7B103AASB6 CRC64;
 Query Match 1.8%; Score 8; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 172 WMAPEVIO 179
 DB 171 WMAPEVIO 178
 RESULT 19
 ID T2EA_YEAST STANDARD; PRT; 482 AA.
 AC P36100;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription initiation factor TFIIF, alpha subunit (TFIIF-alpha)
 DE (Transcription factor A large subunit) (Factor A 66 kDa subunit).
 GN TPA1 OR YKL028W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 66-71, 102-108 AND 206-220.
 RC STRAIN=Bj926;
 RX MEDLINE=95050500; PubMed=7961670;
 RA Feaver W.J., Henry N.L., Bushnell D.A., Sayre M.H., Brickner J.H.,
 RA Gileadi O., Kornberg R.D.;
 RT "Yeast TFIIF. Cloning, expression, and homology to vertebrate
 RT proteins.";
 RL J. Biol. Chem. 269:27549-27553 (1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Rieger M.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECRUITS TFIIF TO THE INITIATION COMPLEX AND STIMULATES
 CC THE RNA POLYMERASE II C-TERMINAL DOMAIN KINASE AND DNA-DEPENDENT
 CC ATPASE ACTIVITIES OF TFIIF. BOTH TFIIF AND TFIIE ARE REQUIRED FOR
 CC PROMOTER CLEARANCE BY RNA POLYMERASE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE TFIIE ALPHA SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U12825; AA62665.1; -
 DR EMBL: Z28028; CA81863.1; -
 DR PIR: S37845; S37845.
 DR TRANSFAC: T02239; -
 DR SGD: S0001511; TPA1.
 DR InterPro: IPR002853; TFIIF_alpha.
 DR Pfam: PR02002; TFIIF_alpha; 1.
 DR SMART: SM00531; TFIIF; 1.
 DR Transcription regulation; Nuclear protein; Zinc-finger.
 FT ZN FING 124 152 C4-TYPE (POTENTIAL).
 FT DOMAIN 374 482 ASP/GLU-RICH (HIGHLY ACIDIC).
 FT POLY-GLU 374 392 POLY-GLU.
 SQ SEQUENCE 482 AA; 54742 MW; 3F789DCA447E8A CRC64;
 Query Match 1.8%; Score 8; DB 1; Length 482;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 436 GDDDDDD 443

```

Db          464 GDDDDDD 471

RESULT 20
STK4_HUMAN STANDARD; PRT; 487 AA.
AC Q130F3; Q15802; Q9VNT4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase
MST1) (MST-1) (Mammalian STE20-like protein kinase 1)
DE (Serine/threonine protein kinase Krs-2).
GN STK4 OR MST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95394929; PubMed=7665586;
RA Creasy C.L., Chernoff J.;
RT "Cloning and characterization of a human protein kinase with homology
to Ste20."
RL J. Biol. Chem. 270:21695-21700(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413604; PubMed=8816758;
RA Taylor L.K., Wang H.C., Erikson R.L.;
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
2."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
RN [3]
RP SEQUENCE OF 1-435 FROM N.A.
RA Laird G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
THE KINASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
-----
DR EMBL: U18297; AAA83254.1; -
DR EMBL: U60207; AAB17262.1; -
DR EMBL: AL109839; CAB89421.1; -
DR HSSP: P24941; IHCL.
DR Genew; HGNC:11408; STK4.
DR MIM; 604965; -
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```

```

KM Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 30 281 PROTEIN KINASE.
FT NP_BIND 36 44 ATP (BY SIMILARITY).
FT BINDING 59 59 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
FT DOMAIN 173 378 POLY-GLU.
FT CONFLICT 222 222 P -> R (IN REF. 1).
FT CONFLICT 312 312 V -> M (IN REF. 1).
SQ SEQUENCE 487 AA; 55630 MW; 150758B5CF77D5C CRC64;

Query Match 1.8%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIQ 179
Db 190 WMAPEVIQ 197

RESULT 21
STK3_HUMAN STANDARD; PRT; 491 AA.
AC Q13188; Q15445; Q15801;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase 3 (EC 2.7.1.37) (STE20-like kinase
MST2) (MST-2) (Mammalian STE20-like protein kinase 2)
DE (Serine/threonine protein kinase Krs-1).
GN STK3 OR MST2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144292; PubMed=8566796;
RA Creasy C.L., Chernoff J.;
RT "Cloning and characterization of a member of the MST subfamily of
Ste20-like kinases."
RL Gene 167:303-306(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413604; PubMed=8816758;
RA Taylor L.K., Wang H.C., Erikson R.L.;
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
2."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
RN [3]
RP SEQUENCE OF 96-203 FROM N.A.
RX MEDLINE=94100173; PubMed=8274451;
RA Schultz S.J., Nigg E.A.;
RT "Identification of 21 novel human protein kinases, including 3 members
of a family related to the cell cycle regulator nimA of Aspergillus
nidulans."
RL Cell Growth Differ. 4:821-830(1993).
RN [4]
RP FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT
MAY PLAY A ROLE IN THE RESPONSES TO ENVIRONMENTAL STRESS (BY
SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,
SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT
HEART, LUNG AND BRAIN TISSUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U26424; AAC50386.1; -
CC EMBL: U60206; AAB17261.1; -
CC EMBL: Z25422; CAA80909.1; -
CC PIR: I38212; I38212.
CC HSSP: P24941; 1HCL.
CC DR GeneW; HGNC:11406; STR3.
CC DR MIM: 605030.
CC DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
CC DR GO; GO:0004668; P:protein amino acid phosphorylation; TAS.
CC DR GO; GO:0007165; P:signal transduction; TAS.
CC DR InterPro; IPR000719; Prot kinase.
CC DR InterPro; IPR002290; Ser Thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC DR PROSITE; PS00108; PROTEIN KINASE ST; FALSE_NEG.
CC DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC KM Transferase; Serine/threonine-protein kinase; ATP-binding.
CC FT DOMAIN; 27 278 PROTEIN KINASE.
CC FT NP_BIND; 33 41 ATP (BY SIMILARITY).
CC FT BINDING; 56 56 ATP (BY SIMILARITY).
CC FT ACT_SITE; 146 146 BY SIMILARITY.
CC FT DOMAIN; 308 314 POLY-GLU.
CC FT DOMAIN; 370 375 POLY-GLU.
CC FT CONFLICT; 96 98 WIV -> YLY (IN REF. 3).
CC FT CONFLICT; 121 121 D -> Y (IN REF. 3).
CC FT CONFLICT; 203 203 D -> G (IN REF. 3).
CC FT CONFLICT; 303 303 D -> E (IN REF. 2).
CC FT CONFLICT; 332 334 GEC -> ESY (IN REF. 2).
CC SQ SEQUENCE 491 AA; 56261 MW; 9CAJB0644F3C1A9 CRC64;

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 491;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
DB 187 WMAPEVIO 194

RESULT 22.
MK7_MOUSE STANDARD; PRT; 579 AA.
ID MK7_MOUSE
AC 062073;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
DE MAP3K7 OR TAK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96123277; PubMed=8533096;
RA Yamauchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
RA Taniguchi T., Nishida E., Matsunoto K.,
RT Identification of a member of the WAPKK family as a potential
RT mediator of TGF-beta signal transduction.";
RL Science 270:2008-2011(1995).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MARKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D76446; BAA11184.1; -
CC HSSP: P08631; 1ADS.
CC DR MGD; MGI:1346877; Map3k7.
CC DR InterPro; IPR000719; Prot kinase.
CC DR InterPro; IPR002290; Ser Thr_pkinase.
CC DR InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC KM Transferase; Serine/threonine-protein kinase; ATP-binding.
CC FT DOMAIN; 8 16 POLY-SER.
CC FT DOMAIN; 36 291 PROTEIN KINASE.
CC FT NP_BIND; 42 50 ATP (BY SIMILARITY).
CC FT BINDING; 63 63 ATP (BY SIMILARITY).
CC FT ACT_SITE; 156 156 BY SIMILARITY.
CC SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 579;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
DB 170 VLKICDFG 177

RESULT 23.
YM72_YEAST STANDARD; PRT; 590 AA.
ID YM72_YEAST
AC 005021;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 67.6 kDa protein in MKPL44-WTFL intergenic region.
DE YMR227C OR YM959.09C.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCB1_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hallin N., Hunt S.,
RA Jagsels R., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skellton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.";
RL Nature 387:90-93(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z49939; CAA90198.1; -
CC PIR: S57594; S57594.
CC TRANSFAC; T03086; -

```

```

DR SGD; S0004840; TAF67.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0000114; P:gl-specific transcription in mitotic cell c. . .; IPI.
DR InterPro; IPR006751; TAF1155_N.1.
DR Pfam; PF04658; TAF1155_N.1.
DR Hypothenical protein; Coiled coil.
FT DOMAIN 50 53 POLY-LYS.
FT DOMAIN 200 203 POLY-GLU.
FT DOMAIN 368 373 POLY-GLU.
FT DOMAIN 413 421 POLY-ASP.
FT DOMAIN 517 531 POLY-GLU.
FT DOMAIN 427 549 COILED COIL (POTENTIAL).
SQ SEQUENCE 590 AA; 67555 MW; C014E7419B0B1C61 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 590;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 413 DDDDDDDG 420

RESULT 24
M3K7_HUMAN STANDARD; PRT; 606 AA.
ID M3K7_HUMAN STANDARD; PRT; 606 AA.
AC 043318; 043317; 043319;
AT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).
RC TISSUE=Lung;
RX MEDLINE=98153801; PubMed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1B;
CC IsoId=043318-1; Sequence=Displayed;
CC Name=1A;
CC IsoId=043318-2; Sequence=VSP_004886;
CC Name=1C;
CC IsoId=043318-3; Sequence=VSP_004887; VSP_004888;
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP
CC KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009357; BAA25026.1; -.
DR EMBL; AB009356; BAA25025.1; -.
DR EMBL; AB009358; BAA25027.2; -.
DR PIR; JCS955; JCS955.

```

```

DR PIR; JCS955; JCS956.
DR HSSP; P08631; 1AD5.
DR Genem; HGNC:6859; MAP3K7.
DR MIM; 602614; -.
DR GO; GO:0004709; F:MAP kinase kinase kinase activity; TAS.
DR GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 8 14 POLY-SER.
FT DOMAIN 36 291 PROTEIN KINASE.
FT NP_BIND 42 50 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARSPPLIC 404 430 Missing (in isoform 1A).
FT VARSPPLIC 509 518 /FTId=VSP_004886.
FT VARSPPLIC 519 606 /FTId=VSP_004887.
FT VARSPPLIC 519 606 Missing (in isoform 1C).
SQ SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
DB 170 VLKICDFG 177

RESULT 25
EZH2_HUMAN STANDARD; PRT; 746 AA.
ID EZH2_HUMAN STANDARD; PRT; 746 AA.
AC 015910; 015755; 092857;
AT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Enhancer of zeste homolog 2 (ENX-1).
GN EZH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97124843; PubMed=8954776;
RA Chen H., Roessler C., Antonarakis S.E.;
RT "Cloning of a human homolog of the Drosophila enhancer of zeste gene
RT (EZH2) that maps to chromosome 21q22.2.";
RL Genomics 38:30-37(1996).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=97357309; PubMed=9214638;
RA Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebersorger A.,
RA Popkin D., Pillus L., Jenuwein T.;
RT "Mammalian homologues of the Polycomb-group gene Enhancer of zeste
RT mediate gene silencing in Drosophila heterochromatin and at S.
RT cerevisiae telomeres.";
RL EMBO J. 16:3219-3232(1997).
RN [3]
SEQUENCE OF 134-746 FROM N.A.
RX MEDLINE=96220494; PubMed=8649418;
RA Hobert O., Talian B., Ulrich A.;

```

```

RT "Interaction of Vav with ENX-1, a putative transcriptional regulator
RT of homeobox gene expression."
RL Mol. Cell. Biol. 16:3066-3073 (1996).
RN (4)
RP ATRX BINDING.
RX MEDLINE=98167853; PubMed=9499421;
RA Cardoso C., Timet S., Villard L., Khreschatisky M., Fontes M.,
RA Colleaux L.;
RT "Specific interaction between the XNP/ATR-X gene product and the SET
RT domain of the human EZH2 protein."
RL Hum. Mol. Genet. 7:679-684 (1998).
CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION
CC AND CHROMATIN STRUCTURE.
CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (PROBABLE).
CC -1- TISSUE SPECIFICITY: Expressed in many tissues.
CC -1- SIMILARITY: Contains 1 SET domain.
CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95653; CAA64955.1; -
CC EMBL; U61145; AAC51520.1; -
CC EMBL; U52965; AAC50591.1; -
CC PIR; G02838; G02838.
CC TRANSFAC; T04888; -
CC Genew; HGNC:3527; EZH2.
CC MIM; 601573; -
CC DR GO; GO:0003677; F:DNA binding activity; TAS.
CC DR GO; GO:0006325; P:establishment and/or maintenance of chromatin; TAS.
CC DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
CC DR InterPro; IPR001005; Myb_DNA_binding.
CC DR InterPro; IPR001214; SET.
CC DR Pfam; PF00856; SET; 1.
CC DR SMART; SM00717; SANT; 2.
CC DR SMART; SM00317; SET; 1.
CC DR PROSITE; PS50280; SET; 1.
CC KW Transcription regulation; Nuclear protein; DNA-binding.
CC FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 523 605 CYS-RICH.
CC FT DOMAIN 611 731 SET.
CC FT DOMAIN 224 224 F -> L (IN REF. 1).
CC FT CONFLICT 724 724 F -> V (IN REF. 1).
CC FT CONFLICT 724 724 F -> V (IN REF. 1).
CC SQ SEQUENCE 746 AA; 85363 MW; 1B5029EB9D509B85 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 746;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 183 DDDDDDDG 190

RESULT 26
EZH2_MOUSE
ID EZH2_MOUSE STANDARD; PRT; 746 AA.
AC Q61188; Q9R090;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enhancer of zeste homolog 2 (ENX-1).
DE EZH2 OR ENX1H.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10990;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97014262; PubMed=8861097;
RA Hobert O., Sures I., Ciosek T., Fuchs M., Ullrich A.;
RT "Isolation and developmental expression analysis of Enx-1, a novel
RT mouse polycomb group gene."
RL Mech. Dev. 55:171-184 (1996).
RN (2)
RP SEQUENCE OF 134-497 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=99160476; PubMed=1005131;
RA Laible G., Haynes A.R., Lebeisgerger A., O'Carroll D., Mattei M.-G.,
RA Denny P., Brown S.D., Denuwein T.;
RT "The murine polycomb-group genes enh1 and enh2 map close to hox gene
RT clusters on mouse chromosomes 11 and 6."
RL Mamm. Genome 10:311-314 (1999).
CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION
CC AND CHROMATIN STRUCTURE.
CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=ENX-1A;
CC IsoId=Q61188-1; Sequence=Displayed;
CC Name=ENX-1B;
CC IsoId=Q61188-2; Sequence=VSP_001501;
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN EARLY EMBRYOS. IN LATER
CC EMBRYOGENESIS, EXPRESSION RESTRICTED TO CENTRAL AND PERIPHERAL
CC NERVOUS SYSTEM, LIVER AND THYMUS. IN ADULT, HIGHEST EXPRESSION IN
CC SPLEEN, TESTIS AND PLACENTA. LOWER LEVELS IN INTESITINE AND MUSCLE
CC AND VERY LOW LEVELS IN BRAIN AND LIVER. NO EXPRESSION IN HEART,
CC THYROID GLAND, LUNG AND KIDNEY.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN BOTH ADULT AND EMBRYO WITH
CC HIGHEST LEVELS IN EARLY EMBRYOGENESIS.
CC -1- SIMILARITY: Contains 1 SET domain.
CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52951; AAC52655.1; -
CC EMBL; AF104359; AAD54020.1; -
CC DR MGD; MGI:107940; Ezh2.
CC DR InterPro; IPR001005; Myb_DNA_binding.
CC DR InterPro; IPR001214; SET.
CC DR Pfam; PF00856; SET; 1.
CC DR SMART; SM00717; SANT; 2.
CC DR SMART; SM00317; SET; 1.
CC DR PROSITE; PS50280; SET; 1.
CC KW Transcription regulation; Nuclear protein; DNA-binding;
CC KW Alternative splicing.
CC FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 523 605 CYS-RICH.
CC FT DOMAIN 611 731 SET.
CC FT VARSPLIC 511 553
CC FT DSSSNHYNYGQCDHPRQCDSSCPVIONFCRKCQCSS
CC EC -> G (in isoform ENX-1B).
CC /FTID=VSP_001501.
CC MISSING (IN REF. 2).
CC SQ SEQUENCE 746 AA; 85336 MW; 0435C021963ED24E CRC64;

Query Match 1.8%; Score 8; DB 1; Length 746;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 183 DDDDDDDG 190

```

```

RESULT 27
CN3B_MOUSE STANDARD; PRT; 799 AA.
AC 061409;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CGMP-inhibited 3',5'-cyclic phosphodiesterase B (EC 3.1.4.17) (Cyclic
GN GMP-inhibited phosphodiesterase B) (CGIPDEB) (Fragment).
PDEB3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss; TISSUE=Fat;
RX MEDLINE=97079687; PubMed=8921398;
RA Loebbert R.W., Winterpacht A., Seipel B., Zabel B.U.;
RT "Molecular cloning and chromosomal assignment of the human homologue
RT of the rat CGMP-inhibited phosphodiesterase 1 (PDE3A) -- a gene
RT involved in fat metabolism located at 11p15.1.";
RL Genomics 37:211-218 (1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN FAT METABOLISM.
CC -1- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =
CC nucleoside 5'-phosphate.
CC -1- ENZYME REGULATION: INHIBITED BY CGMP.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -1- TISSUE SPECIFICITY: ABUNDANT IN ADIPOSE TISSUES.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95521; CAA64775.1; -.
DR MGD: MGI:1333863; Pde3b.
DR InterPro: IPR003607; Met_phosphohydro.
DR InterPro: IPR002073; PDEase.
DR Pfam: PF00233; PDEase; 1.
DR SMART: SM00471; HDC; 1.
DR PROSITE: PS00126; PDEASE_1; 1.
KM Hydrolyase; CGMP; Membrane.
FT NON_TER 1 1
FT DOMAIN 715 724 POLY-ASP.
FT DOMAIN 773 777 POLY-GLU.
FT NON_TER 799 799
SQ SEQUENCE 799 AA; 89811 MW; 876658959F1C8B3F CRC64;

Query Match 1.8%; Score 8; DB 1; Length 799;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 716 DDDDDDDG 723

```

```

GN YHR131C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Pavello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082 (1994).
CC -1- SIMILARITY: TO YEAST YNL144C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10398; AAB68414.1; -.
DR PIR: S48975; S48975.
DR HSSP: 000963; IDRO.
DR SGD: S0001173; YHR131C.
DR InterPro: IPR001849; PH.
DR Pfam: PF00169; PH; 1.
DR SMART: SM00233; PH; 1.
KM Hypothetical Protein.
FT DOMAIN 312 325 POLY-ARG.
FT DOMAIN 788 840 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 840 AA; 95058 MW; 3A86717D3332A0DF CRC64;

Query Match 1.8%; Score 8; DB 1; Length 840;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 829 DDDDDDDG 836

RESULT 29
DD10_HUMAN STANDARD; PRT; 875 AA.
AC 013206;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable ATP-dependent RNA helicase DDX10 (DEAD-box protein 10).
GN DDX10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96301396; PubMed=8660968;
RA Savitsky K., Ziv Y., Bar-Shira A., Gilad S., Tagle D.A., Smith S.,
RA Uziel T., Sfez S., Nahmias J., Sartiel A., Eddy R.L., Shown T.B.,
RA Collins F.S., Shihon Y., Rotman G.;
RT "A human gene (DDX10) encoding a putative DEAD-box RNA helicase at
RL 11q22-q23."
RL Genomics 33:199-206 (1996).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RA Arai Y., Kaneko Y., Kubo T., Arai K., Hosoda F., Ohki M.;
RT "Molecular analysis of the chromosomal breakpoints and identification
of the repetitive sequences near the breakpoints of NUP98 in
therapy-related leukemia with inv(11)(p15q22).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.
CC -1- TISSUE SPECIFICITY: HIGH IN TESTIS BUT WIDELY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DBP4/DDX10
SUBFAMILY.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.infobiogen.fr/services/chromancer/Genes/DDX10.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28042; AAC50823.1; -.
DR EMBL; AB040537; BAB18536.1; -.
DR SWISS-2DPAGE; Q13206; HUMAN.
DR Genew; HGNC:2735; DDX10.
DR MIM; 601235; -.
DR GO; GO:0003724; F:RNA helicase activity; TAS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
DR KMW Helicase; ATP-binding; RNA-binding.
FT NP BIND 113 120 ATP (POTENTIAL).
FT SITE 222 225 DEAD BOX.
FT CONFLICT 647 647 A -> D (IN REF. 2).
FT CONFLICT 658 658 D -> E (IN REF. 2).
FT CONFLICT 661 661 N -> K (IN REF. 2).
SQ SEQUENCE 875 AA; 100815 MW; 4632ED8A56AD945B5 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 875;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 782 DDDDDDDG 789

RESULT 30
CAFA_HUMAN
ID ID CAFA_HUMAN STANDARD; PRT; 938 AA.
AC Q13111; OSUYU8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chromatin assembly factor 1 subunit A (CAF-1 subunit A) (Chromatin
DE assembly factor I p150 subunit) (CAF-I 150 kDa subunit) (CAF-Ip150).
GN CHAF1A OR CAF1P150 OR CAF.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
OX [1]
RN
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cervical adenocarcinoma;
RX MEDLINE=95323966; PubMed=7606578;
RA Kaufman P.D., Kobayashi R., Kessler N., Stillman B.;
RT "The p150 and p60 subunits of chromatin assembly factor I: a molecular
link between newly synthesized histones and DNA replication.";
```

```

RL Cell 81.1105-1114(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
RX MEDLINE=21147928; PubMed=11250073;
RA Dong H., Lin W., Zhang C.-K., Xiong H., Fu G., Jin W.-R., Chen R.,
RA Chen Z., Qi Z.-T., Huang G.M.;
RT "Genomic sequence and expression analyses of human chromatin assembly
RT factor 1 p150 gene.";
RL Gene 264:187-196(2001).
RN [3]
RP INTERACTION WITH PCNA, AND SUBCELLULAR LOCATION.
RX MEDLINE=99159824; PubMed=10052459;
RA Shibahara K., Stillman B.;
RT "Replication-dependent marking of DNA by PCNA facilitates
RT CAF-I-coupled inheritance of chromatin.";
RL Cell 96:575-585(1999).
RN [4]
RP INTERACTION WITH PCNA.
RX MEDLINE=20115866; PubMed=10648606;
RA Moggs J.G., Grandt P., Quivy J.P., Jonsson Z.O., Hubscher U.,
RA Becker P.B., Almouzni G.;
RT "A CAF-I-PCNA-mediated chromatin assembly pathway triggered by sensing
RT DNA damage.";
RL Mol. Cell. Biol. 20:1206-1218(2000).
RN [5]
RP REVIEW.
RX MEDLINE=20351441; PubMed=10893180;
RA Ridgway P., Almouzni G.;
RT "CAF-I and the inheritance of chromatin states: at the crossroads of
RT DNA replication and repair.";
RL J. Cell Sci. 113:2647-2658(2000).
CC -1- FUNCTION: COMPLEX THAT IS THOUGHT TO MEDIATE CHROMATIN ASSEMBLY IN
DNA REPLICATION AND DNA REPAIR. ASSEMBLES HISTONE OCTAMERS ONTO
REPLICATING DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE
NUCLEOSOME ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES
H3 AND H4 TO REPLICATING DNA; HISTONES H2A/H2B CAN BIND TO THIS
CHROMATIN PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE
HISTONE OCTAMER. P150 BINDS TO HISTONES H3 AND H4. IT MAY PLAY A
ROLE IN HETEROCHROMATIN MAINTENANCE IN PROLIFERATING CELLS BY
BRINGING NEWLY SYNTHESIZED CBX PROTEINS TO HETEROCHROMATIN DNA
REPLICATION FOCI (BY SIMILARITY).
CC -1- SUBUNIT: CAF-1 IS COMPOSED OF THREE SUBUNITS, P48, P60 AND P150.
P150 BINDS DIRECTLY TO P60. ONLY MINOR AMOUNTS OF P48 ARE
COMPLEXED WITH P60 AND P150 IN G1 PHASE. P150 BINDS DIRECTLY TO
PCNA AND TO CBX1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; DNA REPLICATION FOCI.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Comment=Experimental confirmation may be lacking for some
isoforms;
isoform1;
Name=1;
IsoId=Q13111-1; Sequence=Displayed;
isoform2;
Name=2;
IsoId=Q13111-2; Sequence=VSP_004149, VSP_004150;
Name=3;
IsoId=Q13111-3; Sequence=VSP_004151;
CC -1- DEVELOPMENTAL STAGE: ACTIVE COMPLEX IS FOUND IN G1, S AND G2
PHASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U20979; AAA76736.1; -.
DR EMBL; AF190465; AAF04291.1; -.
DR PIR; A56731; A56731.
DR Genew; HGNC:1910; CHAF1A.
DR MIM; 601246; -.
DR GO; GO:0005678; C:chromatin assembly complex; TAS.
```

```

DR GO: 0003682; F:chromatin binding activity; TAs.
DR GO: 0003762; F:histone-specific chaperone activity; TAs.
DR GO: 0006335; P:DNA replication dependent nucleosome assembly; TAs.
DR GO: 0006461; P:protein complex assembly; TAs.
KW DNA replication; DNA repair; Cell cycle; Chaperone; Nuclear protein;
KW alternative splicing.
FT DOMAIN 1 31 BINDS TO PCNA.
FT DOMAIN 1 296 BINDS TO CBX1 CHROMO SHADOW DOMAIN.
FT DOMAIN 305 435 ARG/GLU/LYS-RICH.
FT DOMAIN 584 590 POLY-GLU.
FT DOMAIN 601 605 POLY-ASP.
FT DOMAIN 887 893 POLY-GLU.
FT DOMAIN 642 938 BINDS TO P60.
FT VARSPLIC 692 731 CPLETLPAGEQCPVCRRLRNSPQSRKLNKNSCVSTLSC
FT (in isoform 2).
FT (in isoform 2).
FT VARSPLIC 732 938 /FTid=VSP_004149.
FT VARSPLIC 754 926 /FTid=VSP_004150.
FT CONFLICT 757 757 /FTid=VSP_004151.
SQ SEQUENCE 938 AA; 105222 MW; DBD01B9F1A0E28B7 CRC64;
Query Match 1.8%; Score 8; DB 1; Length 938;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 434 SEGDDDD 441
Db 598 SEGDDDD 605
RESULT 31
ROR2 HUMAN STANDARD; PRT; 943 AA.
AC 001974; Q9HAY7; Q9HB61;
ID 16-OCT-2001 (Rel. 40; Created)
DT 15-OCT-2001 (Rel. 40; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
GN ROR2 OR NTRK2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCB1; TaxID=9606;
OX 1
RN SEQUENCE FROM N.A.
RX MEDLINE=93100347; PubMed=1334494;
RA Maslakowski P., Catroli R.D.,
RT "A novel family of cell surface receptors with tyrosine kinase-like
RT domain."
RL J. Biol. Chem. 267:26181-26190 (1992).
RN [2]
RP SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
RX MEDLINE=20164326; PubMed=10700192;
RA Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,
RA Pollitt C., DeChiara T.M., Kumble R.B., Valenzuela D.M.,
RA Yancopoulos G.D., Wilkie A.O.M.,
RT "Dominant mutations in ROR2, encoding an orphan receptor tyrosine
RT kinase, cause brachydactyly type B."
RL Nat. Genet. 24:275-278 (2000).
RN [3]
RX SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.
RX MEDLINE=20442029; PubMed=10986040;
RA Schwabe G.C., Tinschelet S., Buschow C., Meisner P., Wolff G.,
RA Gillissen-Keesbach G., Oldridge M., Wilkie A.O.M., Koeneke R.,
RA Mundlos S.,
RT "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause
RT brachydactyly type B."
RL Am. J. Hum. Genet. 67:822-831 (2000).
RN [4]

```

```

RP VARIANTS RRS CYS-184; TRP-189; TRP-366 AND LYS-620.
RX MEDLINE=20392394; PubMed=10932186;
RA Afzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,
RA Terres-Pereira E., Tulesz B., Murday V.A., Patton M.A.,
RA Wilkie A.O.M., Jeffery S.,
RT "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,
RT is caused by mutation of ROR2."
RL Nat. Genet. 25:419-422 (2000).
RN [5]
RP VARIANT RRS TYR-182.
RX MEDLINE=20392395; PubMed=10932187;
RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.,
RT "Mutation of the gene encoding the ROR2 tyrosine kinase causes
RT autosomal recessive Robinow syndrome."
RL Nat. Genet. 25:423-426 (2000).
RN [6]
RP ERRATUM.
RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.,
RL Nat. Genet. 26:383-383 (2000).
CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED
CC IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE
CC REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
CC EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
CC DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES.
CC -1- DISEASE: Defects in ROR2 are a cause of brachydactyly type B1
CC (BDB1) [MIM:113000]. BDB1 is an autosomal dominant skeletal
CC disorder characterized by hypoplasia/aplasia of distal phalanges
CC and nails. In BDB1 the middle phalanges are short but in addition
CC the terminal phalanges are rudimentary or absent. Both fingers and
CC toes are affected. The thumbs and big toes are usually deformed.
CC -1- DISEASE: Defects in ROR2 are a cause of recessive Robinow syndrome
CC (RRS) [MIM:268310]. RRS is an autosomal disorder characterized by
CC skeletal dysplasia with generalized limb bone shortening,
CC segmental defects of the spine, brachydactyly and a dysmorphic
CC facial appearance.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 kinase domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M97639; AAA60276.1; -
DR EMBL: AF294796; AAG01184.2; -
DR EMBL: AF254747; AAG01184.2; JOINED.
DR EMBL: AF254748; AAG01184.2; JOINED.
DR EMBL: AF254749; AAG01184.2; JOINED.
DR EMBL: AF254750; AAG01184.2; JOINED.
DR EMBL: AF254751; AAG01184.2; JOINED.
DR EMBL: AF254752; AAG01184.2; JOINED.
DR EMBL: AF254753; AAG01184.2; JOINED.
DR EMBL: AF254754; AAG01184.2; JOINED.
DR EMBL: AF279762; AAG31132.1; -
DR EMBL: AF279763; AAG31132.1; JOINED.
DR EMBL: AF279764; AAG31132.1; JOINED.
DR EMBL: AF279765; AAG31132.1; JOINED.
DR EMBL: AF279766; AAG31132.1; JOINED.
DR EMBL: AF279767; AAG31132.1; JOINED.
DR EMBL: AF279768; AAG31132.1; JOINED.
DR EMBL: AF279769; AAG31132.1; JOINED.
DR EMBL: AF279770; AAG31132.1; JOINED.
DR EMBL: AF279771; AAG31132.1; JOINED.
DR PIR: B45082; B45082.

```

DR HSSP; P00747; 1KRN.
 DR Genew; HGNC:10257; ROR2.
 DR MIM; 602337; -.
 DR MIM; 113000; -.
 DR MIM; 268310; -.
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:004714; F: transmembrane receptor protein tyrosine kin. . .; TAS.
 DR GO; GO:0007275; P: development; TAS.
 DR GO; GO:0007165; P: signal transduction; TAS.
 DR InterPro; IPR000024; Fz_domain.
 DR InterPro; IPR007110; Ig_Like.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF01392; Fz; 1.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00069; kinase; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000395; Kringle; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00408; IgC2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS50038; Fz; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR TRANSFAC; Tyrosine-protein kinase; ATP-binding; Receptor;
 KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
 KW Immunoglobulin domain; Developmental protein; Polymorphism;
 KW Disease mutation.
 FT SIGNAL 1 33
 FT CHAIN 34 943
 FT DOMAIN 34 403
 FT TRANSMEM 404 424
 FT DOMAIN 425 943
 FT DOMAIN 425 943
 FT DOMAIN 35 145
 FT DOMAIN 169 303
 FT DOMAIN 316 394
 FT DOMAIN 473 746
 FT DOMAIN 753 782
 FT DOMAIN 784 857
 FT DOMAIN 859 882
 FT NP_BIND 479 487
 FT BINDING 507 507
 FT ACT_SITE 615 615
 FT MOD_RES 646 646
 FT DISULFID 83 135
 FT CARBOHYD 70 70
 FT CARBOHYD 188 188
 FT CARBOHYD 318 318
 FT VARIANT 182 182
 FT VARIANT 184 184
 FT VARIANT 189 189
 FT VARIANT 245 245
 FT VARIANT 366 366
 FT VARIANT 620 620
 FT VARIANT 819 819

POTENTIAL.
 TYROSINE-PROTEIN KINASE TRANSMEMBRANE
 RECEPTOR ROR2.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE C2-TYPE.
 FZ.
 KRINGLE.
 PROTEIN KINASE.
 SER/THR-RICH.
 PRO-RICH.
 SER/THR-RICH.
 ATP (BY SIMILARITY).
 ATP (BY SIMILARITY).
 BY SIMILARITY.
 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 C -> Y (in RRS).
 /FTId=VAR_010911.
 R -> C (in RRS).
 /FTId=VAR_010768.
 R -> W (in RRS).
 /FTId=VAR_010769.
 A -> T.
 /FTId=VAR_010912.
 R -> W (in RRS).
 /FTId=VAR_010770.
 N -> K (in RRS).
 /FTId=VAR_010771.
 V -> I.
 /FTId=VAR_010913.

SO SEQUENCE 943 AA; 104726 MW; DBAC1E4622B5ECA0 CRC64;
 Query Match 1.8%; Score 8; DB 1; Length 943;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 191 SYGVVLME 198
 DB 677 SYGVVLME 684
 RESULT 32
 ROR2_MOUSE
 ID ROR2_MOUSE STANDARD; PRT; 944 AA.
 AC 092138;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor
 DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2)
 DE (MOR2).
 GN ROR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9248426; PubMed=10231392;
 RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
 RA Hata T., Akira S., Matsuda Y., Yamamura H., Oreni H., Minami Y.,
 RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
 RT mRor1, mRor2, during mouse development: implications in development
 RT and function of the nervous system.";
 RL Genes Cells 4:41-56(1999).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=20164325; PubMed=10700181;
 RA DeChiara T.M., Kimble R.B., Poueymlrou W.T., Rojas J., Maslakowski P.,
 RA Valenzuela D.M., Yancopoulos G.D.;
 RT "Ror2, encoding a receptor-like tyrosine kinase, is required for
 RT cartilage and growth plate development.";
 RL Nat. Genet. 24:271-274(2000).
 CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED
 CC IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE
 CC REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 fritzled (FZ) domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB010384; BAA75481.1; -.
 CC HSSP; P00747; SHPG.
 CC MGD; MGI:1347521; Ror2.
 CC GO; GO:0001501; P: skeletal development; IMP.
 CC InterPro; IPR000024; Fz_domain.
 CC InterPro; IPR007110; Ig_Like.
 CC InterPro; IPR003598; Ig_C2.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR000719; Prot_kinase.

```

DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF01392; Fz; 1.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00051; Kring1; 1.
DR Pfam: PF00069; Kring2; 1.
DR PRINTS: PR00018; KRING1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000335; Kring1; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00219; TYKC; 1.
DR PROSITE: PS50038; FZ; 1.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS50021; KRING1; 1.
DR PROSITE: PS50070; KRING2; 1.
DR PROSITE: PS50107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR Transferrase: Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kring1; Phosphorylation;
KW Immunoglobulin domain; Developmental protein.
KW SIGNAL.
FT CHAIN 1 33
FT FT 34 944
FT FT 34 403
FT FT 404 424
FT FT 425 944
FT FT 55 145
FT FT 169 303
FT FT 316 394
FT FT 473 746
FT FT 753 782
FT FT 784 857
FT FT 859 882
FT FT 479 507
FT FT 507 507
FT FT ACT SITE 615 615
FT FT MOD_RES 646 646
FT FT DISULFID 83 135
FT FT CARBOHYD 70 70
FT FT CARBOHYD 188 188
FT FT CARBOHYD 318 318
SQ SEQUENCE 944 AA; 105050 MW; CD2EBC710387A56 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 944;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVLMWE 198
DB 677 SYGVLMWE 684

RESULT 33
VP41_YEAST STANDARD; PRT; 992 AA.
AC P38959; P87334; Q12011;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vacuolar assembly protein VP841 (Vacuolar morphogenesis protein VAM2).
GN VP841 OR VAM2 OR FET2 OR YDR080W OR D446 OR YD8554.13.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A; PubMed=9111041;
RX Nakamura N., Hirata A., Ohsumi Y., Wada Y.;
RT "Vam2/Vps41p and Vam6/Vps39p are components of a protein complex on

```

```

RT the vacuolar membranes and involved in the vacuolar assembly in the
RT Yeast Saccharomyces cerevisiae."
RT J. Biol. Chem. 272:11344-11349(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96093910; PubMed=7483840;
RA Coster F., Joniaux J.-L., Goffeau A.;
RT "Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open
RT reading frames, including TP52, PPH3, RAD55, SED1, PDC2, AFR1, SSS1,
RT SLU7 and a tRNA for arginine."
RT Yeast 11:673-679(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Richards C., Harris D.E., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBD databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97303186; PubMed=9159129;
RA Radisky D.C., Snyder W.B., Emr S.D., Kaplan J.;
RT "Characterization of VPS41, a gene required for vacuolar trafficking
RT and high-affinity iron transport in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 94:5662-5666(1997).
CC -1- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY AND VACUOLAR TRAFFIC.
CC -1- SUBUNIT: INTERACTS WITH VPS39/VAM6.
CC -1- SIMILARITY: BELONGS TO THE VPS41 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000223; BA19071.1; -.
DR EMBL: X82086; CA457607.1; -.
DR EMBL: Z74376; CA98989.1; -.
DR EMBL: Z46796; CA86802.1; -.
DR PIR: S49835; S49835.
DR SGD: S0002487; VPS41.
DR GO: GO:0005774; C:vacuole (sensu Fungi); IDA.
DR GO: GO:0000324; C:vacuolar membrane; IDA.
DR GO: GO:0017112; F:RAB guany- nucleotide exchange factor activity; IPI.
DR GO: GO:0042145; F:homotypic vacuole fusion (non-autophagic); IDA.
DR GO: GO:0015031; P:protein transport; IMP.
DR GO: GO:0007033; P:vacuole organization and biogenesis; IMP.
DR InterPro: IPR000547; Clathrin_repeat.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00637; Clathrin; 1.
DR Pfam: PF00400; WD40; 1.
DR SMART: SM00299; CLH; 1.
DR SMART: SM00320; WD40; 2.
FT DOMAIN 79 94
FT DOMAIN 239 244
FT CONFLICT 424 424 K -> M (IN REF. 2)
SQ SEQUENCE 992 AA; 113411 MW; E1B48B45DA14A005 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 78 GDDDDDD 85

RESULT 34
IFH1_YEAST STANDARD; PRT; 1085 AA.
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)

```



```

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE IFH1 protein (RRP3 protein).
GN IFH1 OR RRP3 OR YLR223C OR L6083.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Chereil I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RT Saccharomyces cerevisiae."
RL Yeast 11:261-270(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,
RA Benes V., Brunecker M., Deltis H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Newkitch U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetle D., Punelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Schafte M., Scherens B., Scholier P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhaesselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambut R., Medler E.,
RA Medler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII."
RL Nature 387:87-90(1997).
CC -1- FUNCTION: CONTROLS THE PRE-RNA PROCESSING MACHINERY IN
CC CONJUNCTION WITH PHL1. COULD CONVERT PHL1 FROM A REPRESSOR
CC TO AN ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z29488; CA82624.1; -
CC EMBL: U19027; AAB67412.1; -
CC PIR: S55352; S55352.
CC SCD: S0004213; IFH1.
CC
CC GO: GO:0006348; P:chromatin silencing at telomere; IMP.
CC
CC GO: GO:0006364; P:RNA processing; IGI.
CC
CC KW Nuclear protein; Transcription regulation.
CC DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).
CC SEQUENCE 1085 AA; 122491 MW; BE1C7DEF06213FE0 CRC64;
SO

```

```

Query Match 1.8%; Score 8; DB 1; Length 1085;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
DB 142 DDDDDDDG 149

```

```

RESULT 35
YAB9_YEAST STANDARD; PRT; 1131 AA.
AC P31350;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 128.5 kDa helicase in ATSL-TPD3 intergenic region.

```

```

GN YAL019W OR YAL001 OR FUN30.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=93209532; PubMed=8458570;
RA Ouellette B.F.F., Clark M.W., Keng T., Storms R.K., Zhong W.W.,
RA Zeng B., Fortin N., Delaney S., Barton A.B., Kaback D.B., Bussey H.;
RT "Sequencing of chromosome I from Saccharomyces cerevisiae: analysis
RT of a 32 kb region between the LTR1 and SPO7 genes."
RL Genome 36:32-42(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94193531; PubMed=8144453;
RA Barton A.B., Kaback D.B.;
RT "Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae:
RT analysis of the genes in the FUN30-MAK16-SPO7 region."
RL J. Bacteriol. 176:1872-1880(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=92221690; PubMed=1561836;
RA Clark M.W., Zhong W.W., Keng T., Storms R.K., Barton A.B.,
RA Kaback D.B., Bussey H.;
RT "Identification of a Saccharomyces cerevisiae homolog of the SNF2
RT transcriptional regulator in the DNA sequence of an 8.6 kb region in
RT the LTR1-CYS1 interval on the left arm of chromosome I."
RL Yeast 8:133-145(1992).
CC -1- FUNCTION: MAY FUNCTION AS A NEGATIVE REGULATORY ELEMENT IN THE
CC PROCESS OF UV DAMAGE REPAIR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L05146; AAC04938.1; -
CC PIR: S22266; S22266.
CC SCD: S0000017; FUN30.
CC
CC GO: GO:0007001; P:chromosome organization and biogenesis (sen. .; IGI.
CC
CC InterPro: IPR001410; DEAD.
CC
CC InterPro: IPR001650; Helicase_C.
CC
CC Pfam: PF00271; Helicase_C; 1.
CC
CC SMART: SM00487; DEXDC; 1.
CC
CC SMART: SM00490; HELIC; 1.
CC
CC KW Hypothetical protein; Nuclear protein; DNA-binding; Helicase;
CC ATP-binding.
CC NP BIND 597 604 ATP (POTENTIAL).
CC SITE 703 706 DECH BOX.
CC SEQUENCE 1131 AA; 128506 MW; 7E61B9A8B3A42ED2 CRC64;
SO

```

```

Query Match 1.8%; Score 8; DB 1; Length 1131;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDD 443
DB 524 GDDDDDD 531

```

```

RESULT 36
YNF4_YEAST

```

```

ID YNF4 YEAST STANDARD; PRT; 1165 AA.
AC P5350;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypochemical 128.1 kDa protein in OMP2-MSG5 intergenic region.
GN YNL054W OR N2461 OR YNL2467W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=5288C / FY1676;
RA MEDLINE=96021608; PubMed=8533472;
RT Bergz P., Daignon F., Crouzet M.;
RT "The sequence of a 44 420 bp fragment located on the left arm of
RT chromosome XIV from Saccharomyces cerevisiae.";
RL Yeast 11:967-974 (1995).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; U12141; AAA9658.1; -
CC EMBL; Z71330; CAA95925.1; -
CC PIR; S62982; S62982.
CC SGD; S0004999; VAC7.
CC DR GO; GO:0005737; C:Cytoplasm; IDA.
CC DR GO; GO:0016021; C:Integral to membrane; IDA.
CC DR GO; GO:0000324; C:Vacuole (sensu Fungi); IDA.
CC DR GO; GO:0030234; F:enzyme regulator activity; IGI.
CC DR GO; GO:0006644; P:phospholipid metabolism; IGI.
CC DR GO; GO:0000011; P:vacuole inheritance; IMP.
CC KM Hypochemical protein.
CC SQ SEQUENCE 1165 AA; 128140 MW; 0414978AA3066CE7 CRC64;
Query Match 1.8%; Score 8; DB 1; Length 1165;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 1085 DDDDDDDG 1092

```

```

RT encoding the dinitrogenase gamma subunit."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0125 (RNPH) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF302049; AAG29821.1; -
CC HAMAP; MF_00460; -; 1.
CC DR InterPro; IPR005346; UPF0125.
CC DR Pfam; PF03658; UPF0125; 1.
CC FT DOMAIN 78 86 POLY-ASP
CC SEQUENCE 86 AA; 9632 MW; EBD84A7315BC25A9 CRC64;
SQ
Query Match 1.5%; Score 7; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDD 443
Db 78 DDDDDDD 84

```

```

RESULT 37
RNPH_AZOVI STANDARD; PRT; 86 AA.
ID RNPH_AZOVI
AC Q9F570;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein tnfH.
GN RNPH.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=DJ;
RA Rubio L.M., Rangaraj P., Roberts G.P., Ludden P.W.;
RT "Cloning and mutational analysis of the Azotobacter vinelandii gene

```

```

RESULT 38
KOAA_ECOLI STANDARD; PRT; 101 AA.
ID KOAA_ECOLI
AC P03052;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TRFB transcriptional repressor protein (Regulatory protein KORA).
GN TRFB OR KORA.
OS Escherichia coli.
OC Plasmid IncP-alpha RK2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=84069779; PubMed=6316262;
RA Bechhofer D.H., Figurski D.H.;
RT "Map location and nucleotide sequence of korA, a key regulatory gene
RT of promiscuous plasmid RK2."
RL Nucleic Acids Res. 11:7453-7469 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86232584; PubMed=3520485;
RA Thomas C.M., Smith C.A.;
RT "The trfB region of broad host range plasmid RK2: the nucleotide
RT sequence reveals incC and key regulatory gene trfB/korA/korD as
RT overlapping genes."
RL Nucleic Acids Res. 14:4453-4469 (1986).
CC -1- FUNCTION: IN CONJUNCTION WITH KORB, INHIBITS THE TRANSCRIPTION OF
CC KILA, TRFA AND KORAB OPERONS. IN CONJUNCTION WITH KORC IS
CC RESPONSIBLE FOR THE NEGATIVE CONTROL OF KILC AND KILE OPERONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03962; CAA27596.1; -
CC Plasmid; Transcription regulation; Repressor; DNA-binding.
CC FT DNA BIND 37 56 H-T-H MOTIF (POTENTIAL).
CC SQ SEQUENCE 101 AA; 11306 MW; 25622FB8F98F21A5 CRC64;

```



```

RL Science 269:496-512 (1995).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99182411; PubMed=10080886;
RA Hennig M., Dale G.E., D'Arcy A., Daniel F., Fischer S., Gray C.P.,
RA Jollison S., Mueller F., Page M.G.P., Pattison P., Oefner C.;
RT "The structure and function of the 6-hydroxymethyl-7,8-dihydropterin
RT pyrophosphokinase from Haemophilus influenzae."
RL J. Mol. Biol. 287:211-219 (1999).
CC -1- CATALYTIC ACTIVITY: ATP + 2-amino-4-hydroxy-6-hydroxymethyl-7,8-
CC dihydropteridine = AMP + 2-amino-7,8-dihydro-4-hydroxy-6-
CC (dihydroxymethyl) pteridine.
CC -1- PATHWAY: Dihydrofolate biosynthesis; first step.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: BELONGS TO THE HPPK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32691; AAC21742.1; -
DR PDB; 1CBK; 01-MAR-00.
DR TIGR; HT0064; -
DR InterPro; IPR000550; Hppk.
DR Pfam; PF01288; HPPK; 1.
DR TIGRPFAM; TIGR01498; folK; 1.
DR PROSITE; PS00794; HPPK; 1.
KW Transferrase; Folate biosynthesis; 3D-structure; Complete proteome.
FT STRAND 2 10
FT HELIX 15 27
FT TURN 28 28
FT TURN 30 31
FT STRAND 32 37
FT STRAND 41 43
FT STRAND 54 63
FT HELIX 67 80
FT TURN 81 82
FT STRAND 85 85
FT TURN 88 89
FT STRAND 92 92
FT STRAND 95 101
FT TURN 102 103
FT STRAND 106 107
FT STRAND 112 113
FT HELIX 116 120
FT HELIX 122 131
FT TURN 133 134
FT STRAND 136 136
FT TURN 138 139
FT STRAND 142 142
FT HELIX 143 146
FT HELIX 147 150
FT TURN 151 152
FT STRAND 156 157
SQ SEQUENCE 160 AA; 16299 MW; 136CD15F8844FDDD CRC64;
Query Match 1.5%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 NERTIP 228
DB 108 NERTIP 114
RESULT 41
UBC2 YEAST STANDARD; PRT; 172 AA.
AC P06104;

```

```

DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
GN UBC2 OR RAD6 OR YGL058W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=85113143; PubMed=3881753;
RA Reynolds P., Weber S., Prakash L.;
RT "Rad6 gene of Saccharomyces cerevisiae encodes a protein containing a
RT tract of 13 consecutive aspartates."
RL Proc. Natl. Acad. Sci. U.S.A. 82:168-172 (1985).
RL (2)
RP SEQUENCE FROM N.A.
RX STRAIN=8288C;
RX MEDLINE=97377993; PubMed=9234674;
RA Feuerhahn M., de Montigny U., Potier S., Souciet J.-L.;
RT "The characterization of two new clusters of duplicated genes
RT suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
RT chromosomes."
RL Yeast 13:861-869 (1997).
RL (3)
RP SEQUENCE OF 77-91, AND FUNCTION.
RX MEDLINE=87315384; PubMed=3306404;
RA Jentich S., McGrath J.P., Varshavsky A.;
RT "The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating
RT enzyme."
RL Nature 329:131-134 (1987).
RL (4)
RP MUTAGENESIS OF CYS-88.
RX MEDLINE=90207263; PubMed=2157209;
RA Sung P., Prakash S., Prakash L.;
RT "Mutation of cysteine-88 in the Saccharomyces cerevisiae RAD6 protein
RT abolishes its ubiquitin-conjugating activity and its various
RT biological functions."
RL Proc. Natl. Acad. Sci. U.S.A. 87:2695-2699 (1990).
RL (5)
RP REQUIREMENT FOR E3.
RX MEDLINE=91293094; PubMed=2065660;
RA Sung P., Berleth E., Pickart C., Prakash S., Prakash L.;
RT "Yeast RAD6 encoded ubiquitin-conjugating enzyme mediates protein
RT degradation dependent on the N-end-recognition E3 enzyme."
RL EMBO J. 10:2187-2193 (1991).
RN (6)
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=98165808; PubMed=9497353;
RA Worthyake D.K., Prakash S., Prakash L., Hill C.P.;
RT "Crystal structure of the Saccharomyces cerevisiae ubiquitin-
RT conjugating enzyme Rad6 at 2.6-A resolution."
RL J. Biol. Chem. 273:6271-6276 (1998).
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS. UBC2 IS ACTIVE ON HISTONES. IT IS REQUIRED FOR
CC POSTREPLICATION REPAIR OF UV-DAMAGED DNA AND SPORULATION. UBC2
CC MEDIATES E3-DEPENDENT UBC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC dihydrophosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- SUBUNIT: Interacts with Rad18.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: THE ACIDIC-TAIL DOMAIN OF RAD6 IS IMPORTANT FOR
CC POLYUBIQUITINATION OF HISTONES, AS WELL AS FOR SPORULATION.
CC MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: K02962; AAA34952.1; -
 DR EMBL: 272580; CAA36761.1; -
 DR PIR: A21906; A21906.
 DR PDB: 1AYZ; 26-AUG-98.
 DR SGD: S0003026; RAD6.
 DR GO: GO:0005837; C:26S proteasome; IPI.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0004840; F:ubiquitin conjugating enzyme activity; IDA.
 DR InterPro: IPR000608; UBO_conjugat.
 DR Pfam: PF00179; UO_con. 1.
 DR ProDom: PD000461; UBO_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00183; UBIQUITIN CONJUGAT_1; 1.
 DR PROSITE: PS50127; UBIQUITIN CONJUGAT_2; 1.
 KM Ubl conjugation pathway, Ligase; DNA repair; Sporulation;
 KM Nuclear protein; Multigene family; 3D-structure.
 FT DOMAIN 150 172
 FT MOD_RES 1 1
 FT BINDING 88 88
 FT MUTAGEN 88 88
 FT HELIX 4 18
 FT TURN 22 23
 FT STRAND 24 29
 FT TURN 30 31
 FT STRAND 32 41
 FT TURN 44 45
 FT TURN 47 50
 FT STRAND 52 58
 FT TURN 61 65
 FT STRAND 69 72
 FT TURN 79 80
 FT STRAND 81 81
 FT TURN 83 84
 FT STRAND 86 87
 FT HELIX 90 92
 FT TURN 93 95
 FT TURN 98 99
 FT HELIX 102 113
 FT TURN 114 114
 FT HELIX 124 132
 FT HELIX 134 152
 FT TURN 153 153
 SQ SEQUENCE 172 AA; 19705 MW; 5F568DC28ABBD60F CRC64;

Query Match 1.5%; Score 7; DB 1; Length 172;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDD 443
 DB 157 DDDDDD 163

RESULT 42
 NUGM BRABR
 ID NUGM BRABR STANDARD; PRT; 173 AA.
 AC P43154;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
 GN MTND6 OR ND6 OR NADH6
 OS Brachyramphus brevirostris (Kittlitz murrelet).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Charadriiformes; Alciidae;
 OC Brachyramphus.

OX NCBI_TaxID=28693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94363783; PubMed=8082208;
 RA Moun T., Willassen N.P., Johansen S.;
 RT "Intragenic rearrangements in the mitochondrial NADH dehydrogenase
 RT subunit 6 gene of vertebrates.";
 RL Curr. Genet. 25:554-557(1994).
 CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X73922; CAA52127.1; -
 DR PIR: S44403; S44403.
 DR InterPro: IPR001457; Oxidored_q3.
 DR Pfam: PF00499; oxidored_q3; 1.
 KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 173 AA; 18334 MW; 307D4AD6AC8C3419 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLGASFV 9
 DB 45 SLGASFV 51

RESULT 43
 NUGM BRAMA
 ID NUGM BRAMA STANDARD; PRT; 173 AA.
 AC P43155;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
 GN MTND6 OR ND6 OR NADH6
 OS Brachyramphus marmoratus (Marbled murrelet).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Charadriiformes; Alciidae;
 OC Brachyramphus.
 OX NCBI_TaxID=28694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94363783; PubMed=8082208;
 RA Moun T., Willassen N.P., Johansen S.;
 RT "Intragenic rearrangements in the mitochondrial NADH dehydrogenase
 RT subunit 6 gene of vertebrates.";
 RL Curr. Genet. 25:554-557(1994).
 CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X73923; CAA52128.1; -
 DR PIR: S44404; S44404.
 DR InterPro: IPR001457; Oxidored_q3.
 DR Pfam: PF00499; oxidored_q3; 1.
 KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 173 AA; 18444 MW; F4065804ADF85499 CRC64;

```

Query Match          1.5%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLGASFV 9
    |||||
Db 45 SLGASFV 51

RESULT 44
VIRR_AGRU
ID VIRR_AGRU STANDARD; PRT; 175 AA.
AC 052278;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE VIRR protein.
GN VIRR.
OS Agrobacterium tumefaciens.
OC plasmid pTiA6NC.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=A6NC;
RA Zhu J., Winans S.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0066 (VIRR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AF242881; AAB91563.1; -.
CC InterPro; IPR001378; UPF0066.
CC Pfam; PF01980; UPF0066; 1.
CC Prodom; PD006705; UPF0066; 1.
CC TRFpFams; TIGR00104; TIGR00104; 1.
CC DR PROSITE; PS01318; UPF0066; 1.
CC KM Plasmid.
CC SQ SEQUENCE 175 AA; 18820 MW; 18AF20BB611B087F CRC64;

Query Match          1.5%; Score 7; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 AGAVMHS 398
    |||||
Db 9 AGAVMHS 15

RESULT 45
NO29_XENIA STANDARD; PRT; 183 AA.
ID NO29_XENIA STANDARD; PRT; 183 AA.
AC 042584;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleoplasmmin-like protein NO29 (NOVA).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97470976; PubMed=9326619;

```

```

RA Zlives R.R.F., Schmidt-Zachmann M.S., Franke W.W.;
RT "Identification of a small, very acidic constitutive nucleolar protein
RT (NO29) as a member of the nucleoplasmmin family.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11387-11392(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; Z85983; CAB06552.1; -.
CC InterPro; IPR004301; Nucleoplasmmin.
CC Pfam; PF03066; Nucleoplasmmin; 1.
CC
CC KM Nuclear protein.
CC FT DOMAIN 127 165 ASP/GLU-RICH (HIGHLY ACIDIC).
CC FT DOMAIN 135 145 POLY-GLU.
CC FT DOMAIN 146 162 POLY-ASP.
CC SQ SEQUENCE 183 AA; 20121 MW; C1F5DE503FCE7E9F CRC64;

Query Match          1.5%; Score 7; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443
    |||||
Db 146 DDDDDDD 152

RESULT 46
YE47_THEMA STANDARD; PRT; 196 AA.
ID YE47_THEMA STANDARD; PRT; 196 AA.
AC 09X1F9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein TM1447.
GN TM1447.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AE001796; AAD36515.1; -.
CC PIR; C72253; C72253.
CC TIGR; TM1447; -.
CC HAMAP; MF_01043; -, 1.

```

```

DR   InterPro; IPR003811; DUF205.
DR   Pfam; PF02660; DUF205; 1.
DR   TIGRPFAM; TIGR00023; TIGR00023; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM      4      23      POTENTIAL.
FT   TRANSMEM      52      74      POTENTIAL.
FT   TRANSMEM      110     132     POTENTIAL.
FT   TRANSMEM      137     156     POTENTIAL.
SQ   SEQUENCE      196 AA; 21342 MW; 5AD53DD5C502AB90 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   84 YASLGS 90
DB   137 YASLGS 143

RESULT 47
HMGT_ONCMY STANDARD; PRT; 204 AA.
ID   HMGT_ONCMY
AC   P07746; Q91200;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   High mobility group-T protein (HMG-T) (HMG-T1) (HMG-1).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxId=6022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85269614; PubMed=4022777;
RA   Pentecost B.T., Wright J.M., Dixon G.H.;
RT   Isolation and sequence of cDNA clones coding for a member of the
RT   family of high mobility group proteins (HMG-T) in trout and analysis
RT   of HMG-T-mRNA s in trout tissues.
RL   Nucleic Acids Res. 13:4871-4888(1985).
RN   [2]
RP   REVISIONS.
RA   Wright J.;
RP   Submitted (MAY-1986) to the EMBL/GenBank/DBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=95045507; PubMed=7957172;
RA   Stros M., Nishikawa S., Dixon G.H.;
RT   "cDNA sequence and structure of a gene encoding trout testis high-
RT   mobility-group-1 protein."
RL   Eur. J. Biochem. 225:581-591(1994).
CC   -1- FUNCTION: BINDS PREFERENTIALLY SINGLE-STRANDED DNA AND UNWINDS
CC   DOUBLE STRANDED DNA.
CC   -1- SUBCELLULAR LOCATION: Nuclear.
CC   -1- SIMILARITY: BELONGS TO THE HMG1/HMG2 PROTEIN FAMILY.
CC   -1- SIMILARITY: Contains 2 HMG box domains.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X02666; CAA26500.1; -
DR   EMBL; L32859; AAS5877.1; -
DR   PIR; T01071; T01071.
DR   HSSP; P07155; 1AAB.
DR   InterPro; IPR000135; Highmobily_12.
DR   InterPro; IPR000910; HMG_12_box.
DR   Pfam; PF00505; HMG_box; 2.

```

```

DR   PRINTS; PR00886; HIGHMOBILITY12.
DR   SMART; SM00398; HMG; 2.
DR   PROSITE; PS00353; HMG_BOX_1; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 2.
KW   Nuclear protein; Chromosomal protein; DNA-binding; Repeat.
FT   DNA_BIND      8      78      HMG BOX 1.
FT   DNA_BIND      94     162     HMG BOX 2.
FT   DOMAIN         181     204     HMG BOX 2.
FT   CONFLICT       22      22      ASP/GLU-RICH (ACIDIC).
FT   CONFLICT       149     149      R -> K (IN REF. 3).
SQ   SEQUENCE      204 AA; 23572 MW; 6A858A144AD669A0 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   437 DDDDDDD 443
DB   181 DDDDDDD 187

RESULT 48
HIT_MOUSE
ID   HIT_MOUSE
AC   007133;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Histone H1c (Testicular H1 histone).
DE   H1FT OR H1T OR H1P3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxId=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Leukocyte;
RX   MEDLINE=94060108; PubMed=8241275;
RA   Drabent B., Bode C., Doenecke D.;
RT   Structure and expression of the mouse testicular H1 histone gene
RT   (H1c).
RL   Biochim. Biophys. Acta 1216:311-313(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CD-1; TISSUE=Testis;
RA   van Wert J., Wright J., Wolfe S.A., Grimes S.R.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC   -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF
CC   NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.
CC   -1- SUBCELLULAR LOCATION: Nuclear.
CC   -1- DEVELOPMENTAL STAGE: THIS HISTONE IS A TESTIS-SPECIFIC H1 VARIANT
CC   THAT APPEARS DURING MEIOSIS IN SPERMATOGENESIS.
CC   -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; L28753; AAB58417.1; -
DR   EMBL; X72805; CAA51325.1; -
DR   EMBL; U06232; AAA18359.1; -
DR   PIR; S43434; S43434.
DR   HSSP; P08287; 1GHC.
DR   MGD; MG1:107502; H1f3.
DR   InterPro; IPR005818; Histone_H1/H5.
DR   InterPro; IPR005819; Histone_H5.
DR   Pfam; PF00538; linker histone; 1.
DR   PRINTS; PR00624; HISTONEHS.
DR   SMART; SM00526; H15; 1.

```

KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;
 KM Acetylation; Spermatogenesis; Testis.
 FT INIT_MET 0 BY SIMILARITY.
 FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
 FT DOMAIN 37 110 GLOBULAR.
 FT CONFLICT 163 163 R -> G (IN REF. 1).
 SQ SEQUENCE 207 AA; 21508 MW; A3C647CA97976C44 CRC64;
 Query Match 1.5%; Score 7; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 415 SKRGKK 421
 DB 21 SKRGKK 27
 RESULT 49
 ID PROA_EUGAN STANDARD; PRT; 207 AA.
 AC 08SL54;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (PEP) (Plasid-
 encoded RNA polymerase alpha subunit) (RNA polymerase alpha subunit).
 GN PROA.
 OS Euglena anabaena.
 OC Chloroplast.
 OC Eukaryote; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI_TaxID=38273;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UTEX 373;
 RC MEDLINE=21851312; PubMed=11861918;
 RX Sheveleva E.V., Giordani N.V., Hallick R.B.;
 RT Identification and comparative analysis of the chloroplast alpha-
 subunit gene of DNA-dependent RNA polymerase from seven Euglena
 species.
 RL Nucleic Acids Res. 30:1247-1254(2002).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 of DNA into RNA using the four ribonucleoside triphosphates as
 substrates.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA)(N).
 CC -1- SUBUNIT: In chloroplasts the minimal PEP RNA polymerase is
 composed of four subunits: alpha, beta, beta', and beta" (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.
 CC -1- CAUTION: The C-terminal domain thought to be required for
 interaction with some regulatory factors is missing from this
 protein.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AY047483; ALU83360.1; -
 DR HAMAP, MF_00059; atypical; 1.
 KW Transference; Transcription; DNA-directed RNA polymerase; Chloroplast.
 SQ SEQUENCE 207 AA; 24508 MW; E89F8A9FCDA22 CRC64;
 Query Match 1.5%; Score 7; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 47 LKIKKE 53
 DB 47 LKIKKE 53

DB 149 LKIKKE 155
 RESULT 50
 ID EF1B_ORYSA STANDARD; PRT; 228 AA.
 AC 040680;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Elongation factor 1-beta (EF-1-beta).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. HayaYuki; TISSUE=anther;
 RX MEDLINE=9413985; PubMed=8107147;
 RA Matsumoto S., Terui Y., Shikong X., Taira H., Ejiri S.I.;
 RT Cloning and characterization of the cDNA encoding rice elongation
 factor 1 beta."
 RL FEBS Lett. 338:103-106(1994).
 CC -1- FUNCTION: EF-1-BETA AND EF-1-BETA' STIMULATE THE EXCHANGE OF
 GDP BOUND TO EF-1-ALPHA TO GTP.
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
 BETA', AND GAMMA.
 CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, D23674; BAA04903.1; -
 DR PIR, S41086; S41086.
 DR Gramene; Q40680; -
 DR InterPro; IPR001326; EF1_BD.
 DR Pfam; PF00736; EF1BD_1.
 DR PROSITE; PS00824; EF1BD_2; 1.
 DR PROSITE; PS00825; EF1BD_2; 1.
 KW Elongation factor; Protein biosynthesis.
 FT INIT_MET 0 BY SIMILARITY.
 SQ SEQUENCE 228 AA; 24731 MW; D3CE96C9A7B4EF8D CRC64;
 Query Match 1.5%; Score 7; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDD 443
 DB 100 DDDDDDD 106
 RESULT 51
 ID EF1B_BETVU STANDARD; PRT; 230 AA.
 AC 081918;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Elongation factor 1-beta (EF-1-beta).
 OS Beta vulgaris (Sugar beet).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
 OX NCBI_TaxID=161934;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV. INZUCHTLINIE KWS VV-D/ZR5; TISSUE=leaf;
 RA Viereck R.;
 RT "Nucleotide sequence of the elongation factor 1-beta from Beta
 RT vulgaris";
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EF-1-BETA AND EF-1-BETA' STIMULATE THE EXCHANGE OF
 CC GDP BOUND TO EF-1-ALPHA TO GTP.
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
 CC BETA', AND GAMMA.
 CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 297067; CAB09803.1; -
 DR PIR: T14552; T14552.
 DR InterPro: IPR001326; EFL_BD.
 DR Pfam: PF00736; EFLBD_1.
 DR PROSITE: PS00824; EFLBD_1; 1.
 DR PROSITE: PS00825; EFLBD_2; 1.
 KM Elongation factor; protein biosynthesis.
 FT INIT MET 0
 FT SEQUENCE 230 AA; 24545 MW; EBF1208BFEC41E80 CRC64;
 SQ
 Query Match 1.5%; Score 7; DB 1; Length 220;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDD 443
 Db 101 DDDDDDD 107
 RESULT 52
 ICP3_HSVIN
 ID ICP3_HSVIN STANDARD; PRT; 245 AA.
 AC P37319;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).
 GN ICP34.5.
 OS Herpes simplex virus (type 1 / strain MGH-10).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 NC NCBI_TaxId=37107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90156494; PubMed=2154589;
 RA Chou J., Roizman B.;
 RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
 RT repeats, is conserved in several limited-passage isolates but not in
 RT strain 17syn+";
 RL J. Virol. 64:1014-1020(1990).
 CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M33700; AAA45791.1; -

KM Repeat; late protein.
 FT DOMAIN 155 172 6 X 3 AA TANDEM REPEATS OF A-T-P.
 SQ SEQUENCE 245 AA; 25981 MW; 55AD50AB1B6C411 CRC64;
 QY
 Db 69 DDDDDDD 75
 Query Match 1.5%; Score 7; DB 1; Length 245;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDD 443
 Db 75 DDDDDDD 81
 RESULT 53
 ICP3_HSV1
 ID ICP3_HSV1 STANDARD; PRT; 248 AA.
 AC P36313;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).
 GN ICP34.5.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 NC NCBI_TaxId=10299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88274327; PubMed=2839594;
 RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
 RA McNab D., Perry L.J., Scott J.E., Taylor P.;
 RT "The complete DNA sequence of the long unique region in the genome of
 RT herpes simplex virus type 1";
 RL J. Gen. Virol. 69:1531-1574(1988).
 RN [2]
 RP REVISIONS.
 RX MEDLINE=92341080; PubMed=1321892;
 RA Dolan A., McKie E., McLean A.R., McGeoch D.J.;
 RT "Status of the ICP34.5 gene in herpes simplex virus type 1 strain
 RT 17.";
 RL J. Gen. Virol. 73:971-973(1992).
 CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X14112; -; NOT_ANNOTATED_CDS.
 DR PIR: JQ1682; JQ1682.
 KW Repeat.
 KM DOMAIN 161 175 5 X 3 AA TANDEM REPEATS OF A-T-P.
 FT SEQUENCE 248 AA; 26185 MW; AD334AF0C9CABA65 CRC64;
 SQ
 Query Match 1.5%; Score 7; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDD 443
 Db 75 DDDDDDD 81
 RESULT 54
 YFO4_SCHPO
 ID YFO4_SCHPO STANDARD; PRT; 248 AA.
 AC 094611;

DT 28-FEB-2003 (Rel. 41, Last Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypochemical protein C1296.04 in chromosome I.
 GN SPAC1296.04.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V.J., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Soutou J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin N., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymptez B.,
 RA Welford I., Welford E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesli D., Hilbert H.,
 RA Borzov K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Beer P., Zimmermann W., Medler H., Mambut R., Purnelle B.,
 RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipkavski G.V., Uesery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL, AL035439; CAB36510.3; -;
 DR GeneDB_Spombe; SPAC1296.04; -;
 DR InterPro; IPR006614; DySF_C.
 DR SMART; SM00694; DySFC; 1.
 KM Hypochemical protein.
 SQ SEQUENCE 248 AA; 28583 MW; 3F753507632F4BCC CRC64;

GN AT2G39795 OR T517.20.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 NC Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H.L.,
 RA Moffatt K.S., Cronin L.A., Shen M., Pat G., Van Aken S., Unayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.V., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:761-768(1999).
 CC -----
 CC
 RL Nature 402:761-768(1999).
 CC -----
 CC
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Becker J.R., Theologis A.;
 RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SGP consortium (Salk/Stanford/PGEC)."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MAM33 FAMILY.
 CC -----
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL, AC003000; AAM1485.1; -;
 DR EMBL, AY062757; AAL32835.1; -;
 DR InterPro; IPR003428; MAM33.
 DR Pfam; PF02330; MAM33; 1.
 KM Hypochemical protein; Mitochondrion; Transit peptide.
 FT TRANSIT 1 250
 FT CHAIN ? 250
 FT HYPOTHETICAL PROTEIN AT2G39795.
 SQ SEQUENCE 250 AA; 28061 MW; EB8884B390FE7D47 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 1.5%; Score 7; DB 1; Length 250;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA Chou J., Roizman B.;
 RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
 RT repeats, is conserved in several limited-passagen isolates but not in
 RT strain 17syn+";
 RL J. Virol. 64:1014-1020(1990).
 CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; M33701; AAA45792.1; -;
 KM Repeat; Late protein.
 FT DOMAIN 162 179 6 X 3 AA TANDEM REPEATS OF A-T-P.
 SQ SEQUENCE 252 AA; 26601 MW; A6ED049FEEAA19A CRC64;

 QY Query Match 1.5%; Score 7; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Db 437 DDDDDDD 443
 76 DDDDDDD 82

 RESULT 57
 ASP PLAFS STANDARD; PRT; 253 AA.
 ID -ASP PLAFS STANDARD; PRT; 253 AA.
 AC P13825;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartic acid-rich protein precursor.
 OS Plasmodium falciparum (isolate fcm17 / Senegal).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxId=5845;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87270765; PubMed=3038111;
 RX Lenstra R., D'Auriol L., Andrieu B., le Bras J., Galibert F.;
 RT "Cloning and sequencing of Plasmodium falciparum DNA fragments
 RT containing repetitive regions potentially coding for histidine-rich
 RT proteins: identification of two overlapping reading frames";
 RL Biochem. Biophys. Res. Commun. 146:368-377(1987).
 CC -1- MISCELLANEOUS: THIS PROTEIN IS CODED ON THE REVERSE STRAND OF AN
 CC HISTIDINE-RICH PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY WITH HUMAN SET PROTEIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; M17028; AAA29620.1; -;
 DR PIR; B29653; B29653.
 DR InterPro; IPR002164; NAP_family.
 DR Pfam; PF00956; NAP; 1.
 KM Malaria; Antigen; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 253 POTENTIAL.
 FT DOMAIN 204 253 ASPARTIC ACID-RICH PROTEIN.
 FT SEQUENCE 253 AA; 30248 MW; 83E585DE07AB8504 CRC64;

QY Query Match 1.5%; Score 7; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Db 437 DDDDDDD 443
 218 DDDDDDD 224

 RESULT 58
 LI39 CAEEL STANDARD; PRT; 253 AA.
 ID LI39 CAEEL STANDARD; PRT; 253 AA.
 AC P34684;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein 11n-39.
 GN LIN-39 OR CEH-15 OR C07H6.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=93327430; PubMed=8101475;
 RA Clark S.G., Chisholm A.D., Horvitz H.R.;
 RT "Control of cell fates in the central body region of C. elegans by
 RT the homeobox gene 11n-39.";
 RL Cell 74:43-55(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93327429; PubMed=8101474;
 RA Wang B.B., Mueller-Immergluck M.M., Austin J., Robinson N.T.,
 RA Chisholm A.D., Kenyon C.;
 RT "A homeotic gene cluster patterns the anteroposterior body axis of C.
 RT elegans.";
 RL Cell 74:29-42(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Macri C., Vaudin M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBI databases.
 CC -1- FUNCTION: CONTROLS THE MIGRATION OF NEUROBLASTS AND THE
 CC SUBSEQUENT MID-BODY REGION-SPECIFIC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY.
 CC "DEFORMED" SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; L19639; AAC37168.1; -;
 DR EMBL; L19248; AAB04137.1; -;
 DR EMBL; AC006605; AAK85445.1; -;
 DR PIR; B40722; B40722.
 DR HSSP; P02833; ISAN.
 DR WormBep; C07H6.7; CE03975.
 DR TRANSFAC; T03369; -;
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEOBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEOBOX_1; 1.
 DR PROSITE; PS00071; HOMEOBOX_2; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.
 FT DOMAIN 20 36 POLY-SER.
 FT SITE 121 129 POLY-ASP.
 FT DNA_BIND 148 153 ANTP-TYPE HEXAPEPTIDE.
 FT DOMAIN 165 224 HOMEBOX.
 FT DOMAIN 228 248 PRO-RICH.
 SQ SEQUENCE 253 AA; 27814 MW; 6B59DE9568B034A1 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGGDDDD 441
 DB 119 EGGDDDD 125

RESULT 59
 T2D7 RAT STANDARD; PRT; 253 AA.
 AC 062860;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription initiation factor TFIID 31 kDa subunit (TAFI1-31)
 DE (TAFI1-32) (TAFI132) (Neutonal cell death related gene in neuron -7)
 DE (DN-7)).
 GN TAF9 OR TAF2G OR TAFI131.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP MEDLINE=97312553; PubMed=9168994;
 RX Aoki T., Koike T., Nakano T., Shibahara K., Nishimura H., Kikuchi H.,
 RA Honjo T.;
 RT "TAFI131 gene is induced upon programmed cell death in
 RT differentiated PC12 cells deprived of NGF.";
 RL Biochem. Biophys. Res. Commun. 234:230-234(1997).
 CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID
 CC (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA
 CC POLYMERASE TRANSCRIPTION. TAFI131 IS A COACTIVATOR FOR THE P53
 CC PROTEIN. ALSO INTERACTS WITH THE ACIDIC TRANSCRIPTION FACTOR
 CC PROTEIN 16 (VP16) AS WELL AS WITH THE GENERAL TRANSCRIPTION FACTOR
 CC TFIIB (BY SIMILARITY).
 CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
 CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs) (BY SIMILARITY). TAFI131
 CC AND P53 BIND TO EACH OTHER VIA AMINO ACID RESIDUES IN THE AMINO-
 CC TERMINAL DOMAIN OF P53 THAT ARE ESSENTIAL FOR TRANSCRIPTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TAF2G FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U40188; AAC53201.1; -.
 DR EMBL; U40188; AAC53201.1; -.
 DR PIR; JCS511; JCS511.
 DR InterPro; IPR004822; Histone core.
 DR InterPro; IPR003162; TFIID-31.
 DR Pfam; PF02291; TFIID-31; 1.
 DR ProDom; PD011023; TFIID-31; 1.
 DR Transcription regulation; Nuclear protein.
 KW DOMAIN 238 251 POLY-ASP.
 FT SEQUENCE 253 AA; 27620 MW; 42045091A9B94378 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
 DB 238 DDDDDDD 244

RESULT 60
 ICP3 HSV1F STANDARD; PRT; 263 AA.
 ID ICP3 HSV1F
 AC P08353;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).
 GN ICP34.5.
 OS Herpes simplex virus (type 1 / strain F).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OC NCBI_TaxID=10304;
 RN [1]
 RP SEQUENCE FROM N.A..
 RX MEDLINE=86115412; PubMed=3003394;
 RA Chou J., Roizman B.;
 RT "The terminal a sequence of the herpes simplex virus genome contains
 RT the promoter of a gene located in the repeat sequences of the L
 RT component.";
 RL J. Virol. 57:629-637(1986).
 RN [2]
 RP REVISIONS.
 RX MEDLINE=90156494; PubMed=2154589;
 RA Chou J., Roizman B.;
 RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
 RT repeats, is conserved in several limited-passage isolates but not in
 RT strain 17syn+.";
 RL J. Virol. 64:1014-1020(1990).
 CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M1240; AAA45794.1; ALT_SEQ.
 DR EMBL; M33699; AAA45790.1; -.
 DR Repeat; Late protein.
 FT DOMAIN 161 190
 SQ SEQUENCE 263 AA; 27533 MW; F5084106A08ACB1 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 263;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
 DB 75 DDDDDDD 81

RESULT 61
 OSTP_CHICK STANDARD; PRT; 264 AA.
 ID OSTP_CHICK
 AC P23498;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Osteopontin precursor (Bone sialoprotein 1).
 GN SPPL.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauvia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91236779; PubMed=2033080;
 RA Castagna P., Bet P., Quarto R., Gennari M., Migliaccio G.,
 RT "CDNA cloning and gene expression of chicken osteopontin. Expression
 of osteopontin mRNA in chondrocytes is enhanced by trypsin treatment
 of cells.";
 RL J. Biol. Chem. 266:9944-9949(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91159433; PubMed=2001376;
 RA Moore M.A., Gotoh Y., Rafidi K., Gerstenfeld L.C.;
 RT "Characterization of a cDNA for chicken osteopontin: expression
 during bone development, osteoblast differentiation, and tissue
 distribution.";
 RL Biochemistry 30:2501-2508(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94192994; PubMed=8144023;
 RA Rafidi K., Simkina I., Johnson E., Moore M.A., Gerstenfeld L.C.;
 RT "Characterization of the chicken osteopontin-encoding gene.";
 RL Gene 140:163-169(1994).
 CC -1- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an
 integral part of the mineralized matrix. Probably important to
 cell-matrix interaction.
 CC -1- FUNCTION: Acts as a cytokine involved in enhancing production of
 interferon-gamma and interleukin-12 and reducing production of
 interleukin-10 and is essential in the pathway that leads to type
 I immunity (By similarity).
 CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.
 CC -1- PTM: N- AND O-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X56772; CAA40091.1; -;
 DR EMBL: M59182; AAA62729.1; -;
 DR EMBL: U01844; AAA18584.1; -;
 DR PIR: I51384; AA0019.
 DR InterPro: IPR002038; Osteopontin.
 DR Pfam: PF00865; Osteopontin; 1.
 DR SMART: SM00017; OSTEO; 1.
 DR PROSITE: PS00884; OSTEOPONTIN; 1.
 KW Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;
 KW Phosphorylation; Signal.
 FT SIGNAL 1 16
 FT CHAIN 1 264
 FT SITE 132 134 OSTEOPONTIN.
 FT CARBOHYD 106 106 CELL ATTACHMENT SITE.
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 104 104 D -> G (IN REF. 2).
 FT CONFLICT 216 216 G -> D (IN REF. 1).
 FT CONFLICT 235 235 R -> A (IN REF. 1).
 FT CONFLICT 240 240 T -> A (IN REF. 1).
 SQ SEQUENCE 264 AA; 29162 MW; CA2F17DB70A92591 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 264;
 Best Local Similarity 100.0%; Fred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDD 443
 DB 99 DDDDDD 105
 RESULT 62
 ID T2D7 HUMAN STANDARD; PRT; 264 AA.
 AC 016554;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription initiation factor TFIIID 31 kDa subunit (TAFII-31)
 DE (TAFII-32) (TAFII32).
 GN TAF9 OR TAF2G OR TAFII31.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95281615; PubMed=7761466;
 RA Lu H., Levine A.J.;
 RT "Human TAFII31 protein is a transcriptional coactivator of the p53
 protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5154-5158(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95320160; PubMed=7597030;
 RA Klemm R.D., Goodrich J.A., Zhou S., Tjian R.;
 RT "Molecular cloning and expression of the 32-kDa subunit of human
 TFIIID reveals interactions with VP16 and TFIIIB that mediate
 transcriptional activation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5788-5792(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95396764; PubMed=7667268;
 RA Hisatake K., Ohta T., Takada R., Guernah M., Horikoshi M.,
 RA Nakatani Y., Roeder R.G.;
 RT "Evolutionary conservation of human
 TATA-binding-polyepptide-associated factors TAFII31 and TAFII80 and
 RT interactions of TAFII80 with other TAFs and with general
 RT transcription factors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8195-8199(1995).
 CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID
 (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA
 CC POLYMERASE TRANSCRIPTION. TAFII31 IS A COACTIVATOR FOR THE P53
 CC PROTEIN. ALSO INTERACTS WITH THE ACIDIC TRANSACTIVATOR VIRAL
 CC PROTEIN 16 (VP16) AS WELL AS WITH THE GENERAL TRANSCRIPTION FACTOR
 CC TFIIIB.
 CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
 CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs). TAFII31 AND P53 BIND TO
 CC EACH OTHER VIA AMINO ACID RESIDUES IN THE AMINO-TERMINAL DOMAIN OF
 CC P53 THAT ARE ESSENTIAL FOR TRANSCRIPTION.
 CC -1- SUBCELLULAR LOCATION: Nuclear
 CC -1- SIMILARITY: BELONGS TO THE TAF2G FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U25112; AAA91318.1; -;
 DR EMBL: U21858; AAC50153.1; -;

```

DR EMBL: U30504; AAA84389.1; -.
DR PIR: I39141; I39141.
DR TRANSFAC: T02113; -.
DR Genew: HGNC:11542; TAF9.
DR MIM: 600822; -.
DR GO: GO:0005669; C:transcription factor TFIIID complex; TAS.
DR GO: GO:0016251; F:general RNA polymerase II transcription fac. .; TAS.
DR GO: GO:0008022; F:protein C-terminus binding activity; TAS.
DR GO: GO:0003713; F:transcription co-activator activity; TAS.
DR GO: GO:0003665; F:transcription from Pol II promoter; TAS.
DR InterPro: IPR004822; Histone_core.
DR InterPro: IPR003162; TFIIID-31.
DR Pfam: PF02291; TFIIID-31; 1.
DR Prodom: PD011023; TFIIID-31; 1.
DR Transcription regulation; Nuclear protein.
DR DOMAIN 250 262 POLY-ASP (ACIDIC).
DR SEQUENCE 264 AA; 28974 MW; 1925AEC65DC84C7 CRC64;
SQ
Query Match 1.5%; Score 7; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443
Db 250 DDDDDDD 256

RESULT 63
EPID_XENLA STANDARD; PRT; 265 AA.
ID EPID_XENLA
AC P29693;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Elongation factor 1-delta (EF-1-delta) (P36).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodidae; Xenopus.
NCBI_TaxId=8355;
OX
RN NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RA MEDLINE=92375699; PubMed=1508694;
RA Morales J., Cormier P., Mulner-Lorillon O., Poulhe R., Belle R.;
RT "Molecular cloning of a new guanine nucleotide-exchange protein, EFL
delta."
RL Nucleic Acids Res. 20:4091-4091(1992).
RN [2]
RP PRELIMINARY SEQUENCE OF 15-27; 112-118; 186-192 AND 226-259.
RA MEDLINE=91331996; PubMed=1869528;
RA Janssen G.M.C., Morales J., Schipper A., Labbes J.C.,
RA Mulner-Lorillon O., Belle R., Moeller W.;
RT "A major substrate of maturation promoting factor identified as
RT elongation factor 1 beta gamma delta in Xenopus laevis."
RL J. Biol. Chem. 266:14885-14888(1991).
CC -1- FUNCTION: EF-1-BETA AND EF-1-DELTA STIMULATE THE EXCHANGE OF
CC GDP BOUND TO EF-1-ALPHA TO GTP.
CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
CC DELTA, AND GAMMA.
CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X66837; CAA47313.1; -.
CC PIR: S26280; S26280.
CC InterPro: IPR001326; EFL_BD.

```

```

DR Pfam: PF00736; EFLBD; 1.
DR PROSITE: PS00824; EFLBD_1; 1.
DR PROSITE: PS00825; EFLBD_2; 1.
DR Elongation factor; Protein biosynthesis.
DR SEQUENCE 265 AA; 29237 MW; 61A1898EC3F9E402 CRC64;
SQ
Query Match 1.5%; Score 7; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443
Db 134 DDDDDDD 140

RESULT 64
A32E_HUMAN STANDARD; PRT; 268 AA.
ID A32E_HUMAN
AC Q9BTT0; OBN154; O8MMW9;
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Acidic leucine-rich nuclear phosphoprotein 32 family member E (LAMP-
DE like protein) (LAMP-L).
GN ANP32E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxId=9606;
OX
RN NCB1
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Jiang M.;
RT "Cloning and characterization of a novel human LAMP-L gene."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pulmonary artery, and Tongue;
RA Oshima A., Takahashi-Fuji A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Maeno Y., Nagai K., Isegai T.;
RT "NEO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datcheoun L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schain J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Inhibits activity of protein phosphatase 2A. Does not
CC inhibit protein phosphatase 1. May play a role in cerebellar
CC development (By similarity).

```

```

CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ANP3 FAMILY
CC -1- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AY057381; AAL25814.1; -
CC EMBL: AK095228; BAC04505.1; -
CC EMBL: AK092572; BAC03942.1; ALT_INIT.
CC EMBL: BC003380; AAH03380.1; -
CC Genew; HGNC:16673; ANP32E.
CC InterPro; IPR001611; LRR.
CC Pfam; PF00560; LRR; 3.
CC Leucine-rich repeat; Repeat; Nuclear protein.
CC REPEAT 41 63 LRR 1.
CC REPEAT 64 86 LRR 2.
CC REPEAT 87 111 LRR 3.
CC REPEAT 115 138 LRR 4.
CC REPEAT 185 186 MISSING (IN REF. 1 AND 2).
CC CONFLICT 185 186 MISSING (IN REF. 1 AND 2).
CC SEQUENCE 268 AA; 30692 MW; 99D74AFB59BF971 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 268;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 442 DDGEED 448
DB 261 DDGEED 267

RESULT 65
HAX1_MOUSE STANDARD; PRT; 280 AA.
ID HAX1_MOUSE STANDARD; PRT; 280 AA.
AC O35387;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE HSI-binding protein (HS1-associating protein X-1) (HAX-1).
GN HSI1 OR HAX1.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe T.; Takeshita H.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DIRECTLY ASSOCIATES WITH HSI, THROUGH BINDING TO ITS N-
CC TERMINAL REGION. MAY FUNCTION IN PROMOTING CELL SURVIVAL. MAY ALSO
CC ASSOCIATE WITH CORTACTIN/EMSI IN NONLYMPHOID CELLS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MAINLY IN MITOCHONDRION, BUT ALSO TO A
CC LESSER EXTENT IN ENDOPLASMIC RETICULUM AND NUCLEAR ENVELOPE (BY
CC SIMILARITY).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: WEAK, TO THE NIP3 AND BCL-2 FAMILY PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF023482; AAB81081.1; -
CC MGD; MGI:1346319; Hsldpl.

```

```

DR GO; GO:0005739; C-mitochondrion; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0005515; F:protein binding activity; IPI.
KM Mitochondrion.
FT DOMAIN 30 44 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 115 280 INVOLVED IN HSI BINDING (BY SIMILARITY).
SQ SEQUENCE 280 AA; 31654 MW; 3123979BA10B45D1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 280;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 30 DDDDDDD 36

RESULT 66
YNV9_CAEBL STANDARD; PRT; 289 AA.
ID YNV9_CAEBL STANDARD; PRT; 289 AA.
AC P34572; Q23407;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein T16H12.9 in chromosome III.
GN T16H12.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Smith A.; Berks M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: Z30662; CAA83142.2; -
CC EMBL: Z47357; CAA83142.2; JOINED.
CC EMBL: Z47357; CAA87428.2; -
CC EMBL: Z30662; CAA87428.2; JOINED.
CC WormDep; T16H12.9; CE28843.
CC Hypothetical protein.
CC DOMAIN 134 145 ASP-RICH.
CC SEQUENCE 289 AA; 32186 MW; 425B5C6107108D7F CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 289;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 134 DDDDDDD 140

RESULT 67
UBC3_YEAST STANDARD; PRT; 295 AA.
ID UBC3_YEAST STANDARD; PRT; 295 AA.
AC P14682;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-34 kDa (EC 6.3.2.19)

```

DE (ubiquitin-protein ligase) (ubiquitin carrier protein) (cell division
 DE control protein 34)
 GN UBC3 OR CDC34 OR DNA6 OR YDR054C OR YD9609.08C OR D4211.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88321694; PubMed=2842867;
 RA Geobl M.G., Yochem J., Jentsch S., McGrath J.P., Varshavsky A.,
 RA Byers B.;
 RT "The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating
 RT enzyme";
 RN Science 241:1331-1335(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96381250; PubMed=8789263;
 RA Brandt P., Ramlow S., Otto B., Bioecker H.;
 RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
 RT of Saccharomyces cerevisiae chromosome IV.";
 RN Yeast 12:85-90(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=8288C / AB972;
 RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. CAPABLE, IN VITRO, TO UBIQUITINATE HISTONE H2A.
 CC -1- FUNCTION: MEDIATES THE INITIATION OF DNA REPLICATION (TRANSITION
 CC OF G1 TO S PHASE IN CELL CYCLE).
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DOMAIN: THE ACIDIC C-TERMINAL EXTENSION IS ESSENTIAL FOR THE CELL
 CC CYCLE FUNCTION.
 CC -1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thioester formation.
 CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M21877; AAA35188.1; -;
 DR EMBL; X84162; CA558970.1; -;
 DR EMBL; Z74350; CA98872.1; -;
 DR EMBL; Z49209; CA89083.1; -;
 DR PIR; A41241; A41241.
 DR HSSP; Q02159; 2UCZ.
 DR SGD; S0002461; CDC34.
 DR GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IPI.
 DR InterPro; IPR000608; UQO_conjugat.
 DR Pfam; PF00179; UQ_con; 1.
 DR ProDom; PD000461; UQO_conjugat; 1.
 DR SMART; SMO0212; UBC3; 1.
 DR PROSITE; PS00183; UBIQUITIN_CONUGAT_1; 1.
 DR PROSITE; PS50127; UBIQUITIN_CONUGAT_2; 1.
 KW Ubi conjugation pathway; Ligase; DNA replication; Nuclear protein;
 KW Cell cycle; Cell division; Multigene family.
 FT BINDING 95 95 UBIQUITIN (BY SIMILARITY).
 FT DOMAIN 191 289 ASP/GLU-RICH (ACIDIC).
 SO SEQUENCE 295 AA; 34064 MW; 1CE330C3AB1436DC CRC64;

QY 437 DDDDDDD 443
 DB 255 DDDDDD 261

RESULT 68
 PEX5_CANAL
 ID PEX5_CANAL STANDARD; PRT; 296 AA.
 AC 074711;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peroxisomal targeting signal receptor (Peroxisomal protein PEX10)
 DE (Peroxin-5) (Pxi1 receptor) (Fragment).
 GN PEX5.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxId=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 64385 / 1001;
 RX MEDLINE=98451822; PubMed=9778800;
 RA Navarro-Garcia F., Perez-Diaz R., Negrodo A., Pia J., Nombela C.;
 RT Cloning and sequence of a 3.835 kbp DNA fragment containing the H184
 RT gene and a fragment of a PEX5-like gene from Candida albicans.";
 RL Yeast 14:1147-1157(1998).
 CC -1- FUNCTION: BINDS TO THE C-TERMINAL PTX1-TYPE TRIPEPTIDE PEROXISOMAL
 CC TARGETING SIGNAL (SKL-TYPE) AND PLAYS AN ESSENTIAL ROLE IN
 CC PEROXISOMAL PROTEIN IMPORT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ITS DISTRIBUTION APPEARS TO BE DYNAMIC. IT
 CC IS PROBABLY A CYCLING RECEPTOR FOUND MAINLY IN THE CYTOPLASM AND
 CC AS WELL ASSOCIATED TO THE PEROXISOMAL MEMBRANE THROUGH A DOCKING
 CC FACTOR (PEX13) (BY SIMILARITY).
 CC -1- SIMILARITY: Contains 8 TPR repeats.
 CC -1- SIMILARITY: STRONG, TO OTHER PEROXISOMAL TARGETING SIGNAL
 CC RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ003115; CA05870.1; -;
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 4.
 DR SMART; SMO0028; TPR; 4.
 KW Peroxisome; Repeat; TPR repeat; Transport; Protein transport.
 FT NON_TER 1 1
 FT REPEAT 1
 FT REPEAT <1 33 TPR 2.
 FT REPEAT 34 67 TPR 3.
 FT REPEAT 68 105 TPR 4.
 FT REPEAT 106 143 TPR 5.
 FT REPEAT 144 177 TPR 6.
 FT REPEAT 178 211 TPR 7.
 FT REPEAT 212 245 TPR 8.
 FT REPEAT 245 296 TPR 8.
 SO SEQUENCE 296 AA; 33220 MW; 1F78F4AE76395A0B CRC64;

QY 95 NSNRSEE 101
 DB 190 NSNRSEE 196

Query Match 1.5%; Score 7; DB 1; Length 296;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 69
 MURB_BACSU
 ID MURB_BACSU STANDARD; PRT; 303 AA.

AC P18579; P16669; P37581;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-N-acetylpyruvate decarboxylase (EC 1.1.1.158) (UDP-N-
 acetylmutaric dehydrogenase).
 GN MURB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90078133; PubMed=2556375;
 RT "Nucleotide sequence and insertional inactivation of a Bacillus
 RT subtilis gene that affects cell division, sporulation, and
 RT temperature sensitivity".
 RL J. Bacteriol. 171:6821-6834(1989).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.U., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Enright S.D., Emmerson P.T.,
 RA Enright K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Frits C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grand G.,
 RA Giuseppe G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hooson S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Katsanara Y., Klaerr-Bianchini M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogsh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Solio B.,
 RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzsaecker T.,
 RA Winters P., Wiput A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).
 RN (3)
 RP SEQUENCE OF 1-23 FROM N.A.
 RX MEDLINE=92380484; PubMed=1387377;
 RA Miyao A., Yoshimura A., Sato T., Yamamoto T., Theeragool G.,
 RA Kobayashi Y.,
 RT "Sequence of the Bacillus subtilis homolog of the Escherichia coli
 RT cell-division gene murG".
 RL Gene 118:147-148(1992).
 RN (4)
 RP SEQUENCE OF 250-303 FROM N.A.
 RX MEDLINE=90078134; PubMed=2556376;
 RA Harry E.J., Wake R.G.,
 RT "Cloning and expression of a Bacillus subtilis division initiation
 RT gene for which a homolog has not been identified in another
 RT organism".
 RL J. Bacteriol. 171:6835-6839(1989).
 RN (5)
 RP FUNCTION.
 RX MEDLINE=96060848; PubMed=7590298;
 RA Rowland S.L., Errington J., Wake R.G.,

RT "The Bacillus subtilis cell-division 135-137 degrees region contains
 RT an essential orf with significant similarity to murB and a
 RT dispensable sbp gene".
 RL Gene 164:113-116(1995).
 CC -1- FUNCTION: CELL WALL FORMATION.
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetylmuramate + NADP(+) = UDP-N-
 CC acetyl-3-O-(1-carboxyvinyl)-D-glucosamine + NADPH.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: Peptidoglycan biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the murB family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, M31827; AAA83969.1; -
 DR EMBL, Z59111; CAB13396.1; -
 DR EMBL, D10602; BAA01455.1; -
 DR EMBL, M31800; AAA22392.1; -
 DR PIR, S26500; A43727.
 DR Subtilist, BG10228; murB.
 DR HAMAP, MF_00037; -; 1.
 DR InterPro, IPR003170; MurB.
 DR InterPro, IPR006094; Oxid_FAD_bind.
 DR Pfam, PF01565; FAD_binding_4; 1.
 DR Pfam, PF02673; MurB_C; 1.
 DR TIGRfam, TIGR00179; murB; 1.
 KW Peptidoglycan synthesis; Cell wall; Cell division; Oxidoreductase;
 KW NADP; Flavoprotein; FAD; Complete proteome.
 SQ SEQUENCE 303 AA; 32808 MW; C33F5367C84E8A6E CRC64;
 Qy 306 QELKERE 312
 Db 6 QELKERE 12
 Query Match 1.5%; Score 7; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 70
 CDK3 HUMAN STANDARD; PRT; 305 AA.
 ID Q00576;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cell division protein kinase 3 (EC 2.7.1.1-).
 GN CDK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=92347325; PubMed=1639063;
 RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorke C., Nelson C.,
 RA Harlow E., Tsai L.-H.,
 RT "A family of human cdc2-related protein kinases".
 RL EMBO J. 11:2909-2917(1992).
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
 CC INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE
 CC HISTONE H1.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC2/CDK3 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X66357; CAA47001.1; --
 CC PIR: S23382; S23382.
 CC PDB: 1LEN; 24-APR-02.
 CC Genew: HGNC:1772; CDK3.
 CC GK: Q00526; --
 CC MIM: 123828; --
 CC DR GO: GO:0004693; P: cyclin-dependent protein kinase activity; TAS.
 CC DR GO: GO:0007048; P: oncogenesis; TAS.
 CC DR GO: GO:0000074; P: regulation of cell cycle; TAS.
 CC DR InterPro: IPR000719; Prot. kinase.
 CC DR Pfam: PF00069; pkinase; 1.
 CC DR Prodom: PD000001; Prot. kinase; 1.
 CC DR SMART: SM00220; S_TKC_1.
 CC DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 CC DR PROSITE: PS50011; PROTEIN KINASE DM; 1.
 CC KM Transference: Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 KM Cell division; Mitosis; Phosphorylation; 3D-structure.
 KM DOMAIN 4 286 PROTEIN KINASE.
 FT NE_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 127 127 BY SIMILARITY.
 FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 305 AA; 35045 MW; 1128BE0096E262A CRC64;

Query Match 1.5%; Score 7; DB 1; Length 305;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLK 135
 DB 123 VHRDLK 129

RESULT 71
 MLF_DROME STANDARD; PRT; 309 AA.
 ID MLF_DROME STANDARD; PRT; 309 AA.
 AC Q9NRY0; Q9V7G3; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myeloid leukemia factor (Myelodysplasia-myeloid leukemia factor)
 DE (DMLF).
 GN MLF OR CG8295.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyridroidea; Drosophilidae; Drosophila.
 OC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RX MEDLINE=2057896; Pubmed=11137299;
 RA Ohno K., Takahashi Y., Hirose F., Inoue Y.H., Taguchi O., Nishida Y.,
 RA Matsukage A., Yamaguchi M.;
 RT "Characterization of a Drosophila homologue of the human
 RT myelodysplasia/myeloid leukemia factor (MLF).";
 RL Gene 260:133-143(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkelley;
 RX MEDLINE=20196006; Pubmed=10721112;
 RA Adams M.D., Ceolniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baker E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abrial J.F., Agayani A., An H.-J., Andrews-Plannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Galbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jostali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modyarty C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS.
 RC STRAIN=Berkelley;
 RX MEDLINE=22426069; Pubmed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Ceolniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Interacts with DRE-binding
 CC factor (DREF).
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in unfertilized eggs,
 CC early embryos, pupae and adult males while a low level expression
 CC is found in adult females and larvae.
 CC -1- DEVELOPMENTAL STAGE: High levels are seen in unfertilized eggs and
 CC expression increases slightly during early embryo stages (2-3
 CC hrs). Levels are high in embryos until 4 hrs after fertilization
 CC and then decrease gradually through embryonic and larval stages.
 CC -1- SIMILARITY: BELONGS TO THE MLF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AB043986; BA96391.1; --
 CC DR EMBL: AB003809; AAF58093.2; --
 CC DR FlyBase: FBgn0034051; MLF.
 FT DOMAIN 96 202 INTERACTION WITH DREF.

```

FT DOMAIN 236 243 POLY-ASP.
SQ SEQUENCE 309 AA; 34413 MW; 4C50951D548FD9AD CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 309;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 236 DDDDDDD 242

RESULT 72
DOS2_YEAST STANDARD; PRT; 310 AA.
ID DOS2_YEAST
AC P54858;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DOS2 protein.
GN DOS2 OR YDR068W OR YD9609.22 OR YD8554.01 OR D4267.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Singer J.D., Manning B.E., Formosa T.;
RT "Control of single-copy DNA replication requires genes that act in
RT ubiquitin metabolism.";
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96381250; PubMed=8789263;
RA Brandt P., Ramlow S., Otto B., Bloecker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
RT of Saccharomyces cerevisiae chromosome IV.";
RL Yeast 12:85-90(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTS IN UBIQUITIN METABOLISM AND IS NECESSARY FOR THE
CC CONTROL OF SINGLE-COPY DNA REPLICATION.
CC -1- SIMILARITY: Contains 1 BSD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19857; AAA66522.1; -
DR EMBL: X84162; CAAS8984.1; -
DR EMBL: Z74364; CAA98886.1; -
DR EMBL: Z49209; CAA89097.1; -
DR EMBL: Z46796; CAA86790.1; -
DR PIR: S54052; S54052.
DR SGD; S0002475; DOS2.
DR InterPro; IPR005607; BSD.
DR Pfam; PF03909; BSD; 1.
DR PROSITE; PS50858; BSD; 1.
FT DOMAIN 176 228 BSD.
FT DOMAIN 255 259 BSD.
FT DOMAIN 302 308 POLY-ASP.
SQ SEQUENCE 310 AA; 35964 MW; B9B318093B984713 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 310;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 437 DDDDDDD 443
DB 302 DDDDDDD 308

RESULT 73
MK03_MOUSE STANDARD; PRT; 314 AA.
ID MK03_MOUSE
AC Q63844; Q61531;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.-) (Extracellular signal-
DE regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase) (MAP
DE kinase 1) (MAPK 1) (P44-ERK1) (ERK2) (P44-MAPK) (Microtubule-
DE associated protein-2 kinase) (MNK1) (Fragments).
GN MAPK3 OR PRKM3 OR ERK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-10 FROM N.A.
RC TISSUE=Pre-B cell;
RX MEDLINE=92020947; PubMed=1717989;
RA Crews C.M., Alessandrini A.A., Erikson R.L.;
RT "Mouse Erk-1 gene product is a serine/threonine protein kinase that
RT has the potential to phosphorylate tyrosine.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8845-8849(1991).
RN [2]
RP SEQUENCE OF 11-314 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=91369479; PubMed=1716439;
RA de Miguel C., Kligman D., Patel J., Deterra-Madleigh S.D.;
RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
RT from mouse and rat brain.";
RL DNA Cell Biol. 10:505-514(1991).
RN [3]
RP SEQUENCE OF 105-143 FROM N.A.
RC STRAIN=CBA; TISSUE=Bone marrow;
RX MEDLINE=93185941; PubMed=8444355;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Novel CDC2-related protein kinases produced in murine hematopoietic
RT stem cells.";
RL Gene 124:305-306(1993).
RN [4]
RP SEQUENCE OF 105-139 FROM N.A.
RX MEDLINE=93092802; PubMed=1459009;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Identification of new protein kinase genes, similar to kinases of
RT the cdc2 family and expressed in murine hematopoietic stem cells.";
RL Dokl. Akad. Nauk SSSR 324:893-897(1992).
CC -1- FUNCTION: PHOSPHORYLATES MICROTUBULE-ASSOCIATED PROTEIN-2 (MAP2).
CC MYELIN BASIC PROTEIN (MBP), AND ELK-1; MAY PROMOTE ENTRY IN THE
CC CELL CYCLE.
CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation.
CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S58470; AAB19973.1; -
DR EMBL: X64605; CAA45889.1; -
DR PIR: S28184; S28184.

```

```

DR  HSSP; P27703; 2ERR.
DR  MGD; MGT:1346859; MapK3.
DR  GO; GO:0005137; Cytoplasm; IDA.
DR  GO; GO:0009687; P:organelles; IDA.
DR  InterPro; IPR003527; MAP kin.
DR  InterPro; IPR000719; Prot kinase.
DR  InterPro; IPR002290; Ser Thr kinase.
DR  ProDom; PD000001; Prot kinase; 1.
DR  SMART; SM00220; S_TKc; 1.
DR  PROSITE; PS01351; MAPK; 1.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; PARTIAL.
DR  PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR  PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR  Transferase; Serine/threonine-protein kinase; ATP-binding; cell cycle;
KW  phosphorylation.
FT  NON_TER 1 1 POLY-GLY.
FT  DOMAIN 1 4
FT  NON_CONS 10 11
FT  MOD_RES 137 137 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT  MOD_RES 137 137 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT  MOD_RES 139 139 (PROBABLE).
FT  CONFLICT 112 112 T->P (IN REF. 3 AND 4).
FT  SEQUENCE 314 AA; 36198 MW; F06085E047BF7C22 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 LKICDFG 153
DB 115 LKICDFG 121

RESULT 74
KRAF MSV36
ID KRAF MSV36 STANDARD; PRT; 323 AA.
AC P00512;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase transforming protein raf (EC 2.7.1.-).
GN V-RAF.
OS Murine sarcoma virus 3611.
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxId=11812;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84121298; PubMed=6320371;
RA "A common onc gene sequence transduced by avian carcinoma virus WH2
RT and by murine sarcoma virus 3611."
RL Science 223:813-816(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84172180; PubMed=6324342;
RA Mark G.E., Rapp U.R.;
RT "Primary structure of v-raf: relatedness to the src family of
RT oncogenes."
RL Science 224:285-289(1984).
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-RAF
CC POLYPEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MIL/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR  EMBL; K01691; AAA46579.1; ALT_INIT.
DR  PIR; A00638; TWNV6.
DR  HSSP; P12931; IFMK.
DR  InterPro; IPR000719; Prot kinase.
DR  InterPro; IPR002290; Ser Thr kinase.
DR  Pfam; PF00069; pkinaase; 1.
DR  ProDom; PD000001; Prot kinase; 1.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR  PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR  PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR  PolyProtein; Serine/threonine-protein kinase; Transferase; Oncogene;
KW  ATP-binding.
FT  DOMAIN 24 284 PROTEIN KINASE.
FT  NP_BIND 30 38 ATP (BY SIMILARITY).
FT  BINDING 50 50 ATP (BY SIMILARITY).
FT  ACT_SITE 143 143 BY SIMILARITY.
FT  SEQUENCE 323 AA; 36883 MW; 52A5423A6E362F3 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178
DB 186 WMAPEVI 192

RESULT 75
LUKE STAU
ID LUKE STAU STANDARD; PRT; 323 AA.
AC P31715;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukocidin F subunit precursor (Gamma-hemolysin, H-gamma-I subunit).
GN LUKF.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-66.
RX STRAIN=MRSA NO. 4;
RX MEDLINE=92246945; PubMed=1575738;
RA Rahman A., Nariya H., Izaki K., Kato I., Kamio Y.;
RT "Molecular cloning and nucleotide sequence of leukocidin F-component
RT gene (lukF) from methicillin resistant Staphylococcus aureus."
RL Biochem. Biophys. Res. Commun. 184:640-646(1992).
RN [2]
RP SEQUENCE OF 27-85 AND 322-323.
RX STRAIN=RIMD 310925;
RX MEDLINE=93223836; PubMed=8467905;
RA Kamio Y., Rahman A., Nariya H., Ozawa T., Izaki K.;
RT "The two Staphylococcal bi-component toxins, leukocidin and gamma-
RT hemolysin, share one component in common."
RL FEBS Lett. 321:15-18(1993).
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUCOCYTES. GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S;
CC GAMMA-HEMOLYSIN CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I=F
CC AND H-GAMMA-II).
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; S65052; AAC60446.1; -.
DR  PIR; J01530; J01530.

```

DR InterPro; IPR005830; Aer hem leuk.
 DR InterPro; IPR001340; HemLysn_pore.
 DR InterPro; IPR003963; Staph_bicn_txn.
 DR Pfam; PF01117; Aerolysin_1.
 DR PRINTS; PR01468; BICOMPNTOXIN.
 DR TIGRFAMs; TIGR01002; hlyII; 1.
 KW Hemolysis; Toxin; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 323 LEUKOCIDIN F SUBUNIT.
 SQ SEQUENCE 323 AA; 36567 MW; A5951CF2AAB6C7E8 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 323;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 LSVLSHR 63
 |||||
 DB 250 LSVLSHR 256

Search completed: December 5, 2003, 09:30:47
 Job time : 20 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:27:37 ; Search time 40 Seconds
(without alignments)
2935.348 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSLSGAFVQIKFDLQFFE.....GDDDDDDGREDNDNDNSE 455

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

SPTREMBL 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455	100.0	455	4	Q9HCC4
2	331	72.7	800	4	Q9NYE9
3	331	72.7	800	4	Q9NYL2
4	331	72.7	800	4	Q9HCC5
5	331	72.7	800	4	Q9HDD2
6	166	36.5	454	11	Q9ESL3
7	166	36.5	802	11	Q9ESL4
8	104	22.9	289	11	Q8BR73
9	49	10.8	371	13	Q90ZY8
10	25	5.5	109	11	Q8BUD3
11	11	2.4	855	5	Q01700
12	11	2.4	977	5	Q9VW24
13	9	2.0	141	10	Q9SM42
14	9	2.0	205	5	Q8ST12
15	9	2.0	233	4	O15318
16	9	2.0	297	12	Q98182

17	9	2.0	321	10	Q8GRY7	Q8GRY7 locus japon
18	9	2.0	408	10	Q39585	Q39585 chlamydomon
19	9	2.0	453	10	Q85SD1	Q85SD1 oryza sativ
20	9	2.0	460	10	Q85SG9	Q85SG9 oryza sativ
21	9	2.0	502	11	Q91YV1	Q91YV1 mus musculu
22	9	2.0	502	11	Q91YR0	Q91YR0 mus musculu
23	9	2.0	503	11	Q63559	Q63559 rattus norv
24	9	2.0	513	5	Q8MYF3	Q8MYF3 dictyostrel
25	9	2.0	607	10	Q9PG45	Q9PG45 arabidopsis
26	9	2.0	611	10	Q91KBA	Q91KBA arabidopsis
27	9	2.0	616	5	Q9GRH2	Q9GRH2 sycon rapa
28	9	2.0	652	16	Q9A0J5	Q9A0J5 streptococc
29	9	2.0	652	16	Q8P1L1	Q8P1L1 streptococc
30	9	2.0	652	16	Q8K831	Q8K831 streptococc
31	9	2.0	652	16	Q8ESW1	Q8ESW1 streptococc
32	9	2.0	652	16	Q8E084	Q8E084 streptococc
33	9	2.0	764	3	Q12500	Q12500 saccharomyc
34	9	2.0	789	2	Q9XBP8	Q9XBP8 myxococcus
35	9	2.0	840	11	Q8K062	Q8K062 mus musculu
36	9	2.0	856	10	Q9ZSM8	Q9ZSM8 arabidopsis
37	9	2.0	898	10	Q04246	Q04246 arabidopsis
38	9	2.0	1059	3	Q9EPG2	Q9EPG2 candida alb
39	9	2.0	1472	5	Q905A8	Q905A8 bombyx mori
40	9	2.0	1738	5	Q8IAL5	Q8IAL5 plasmodium
41	9	2.0	1978	5	Q8T5H2	Q8T5H2 anopheles g
42	9	2.0	2391	5	Q27732	Q27732 plasmodium
43	9	1.8	70	10	Q8GVR6	Q8GVR6 oryza sativ
44	8	1.8	101	4	Q15452	Q15452 homo sapien
45	8	1.8	106	12	Q90872	Q90872 spodiopetra
46	8	1.8	145	16	Q8E280	Q8E280 leptospira
47	8	1.8	148	10	Q8LIE2	Q8LIE2 oryza sativ
48	8	1.8	157	12	Q55500	Q55500 human cytom
49	8	1.8	158	5	Q81PM5	Q81PM5 plasmodium
50	8	1.8	161	12	Q919Q2	Q919Q2 culix nigri
51	8	1.8	161	12	Q90867	Q90867 spodiopetra
52	8	1.8	162	16	Q9HMQ2	Q9HMQ2 pseudomonas
53	8	1.8	180	3	Q9P840	Q9P840 candida alb
54	8	1.8	183	4	Q00193	Q00193 homo sapien
55	8	1.8	184	10	Q42341	Q42341 arabidopsis
56	8	1.8	194	12	Q9QAM0	Q9QAM0 white spot
57	8	1.8	194	12	Q91LE7	Q91LE7 white spot
58	8	1.8	194	12	Q8QTE0	Q8QTE0 white spot
59	8	1.8	198	5	Q81P55	Q81P55 drosophila
60	8	1.8	198	10	Q9C7Y9	Q9C7Y9 arabidopsis
61	8	1.8	204	10	Q3756	Q3756 fagus sylvia
62	8	1.8	206	10	Q9SWT7	Q9SWT7 hordeum vul
63	8	1.8	209	10	Q945P0	Q945P0 nicotiana t
64	8	1.8	211	17	Q28960	Q28960 archaeoglob
65	8	1.8	213	11	Q91VH4	Q91VH4 mus musculu
66	8	1.8	218	10	Q8BZK9	Q8BZK9 oryza sativ
67	8	1.8	229	16	Q8X4P6	Q8X4P6 escherichia
68	8	1.8	230	11	Q8C215	Q8C215 mus musculu
69	8	1.8	237	2	Q50439	Q50439 mycobacteri
70	8	1.8	246	10	Q9AS90	Q9AS90 oryza sativ
71	8	1.8	250	10	Q9M1H5	Q9M1H5 arabidopsis
72	8	1.8	258	5	Q9GRQ9	Q9GRQ9 leishmania
73	8	1.8	259	11	Q9CXN1	Q9CXN1 mus musculu
74	8	1.8	260	10	Q9S1S0	Q9S1S0 arabidopsis
75	8	1.8	269	5	Q819F4	Q819F4 plasmodium
76	8	1.8	277	10	Q948F2	Q948F2 oryza sativ
77	8	1.8	277	10	Q8H896	Q8H896 oryza sativ
78	8	1.8	279	11	Q64125	Q64125 mus sp. ear
79	8	1.8	292	4	Q02445	Q02445 helodactyla
80	8	1.8	293	5	Q9NWX4	Q9NWX4 homo sapien
81	8	1.8	293	5	Q81BF3	Q81BF3 plasmodium
82	8	1.8	297	5	Q17319	Q17319 caenorhabdi
83	8	1.8	298	10	Q64936	Q64936 arabidopsis
84	8	1.8	302	12	Q8WB23	Q8WB23 white spot
85	8	1.8	307	5	Q9NKK3	Q9NKK3 drosophila
86	8	1.8	317	11	P97827	P97827 rattus norv
87	8	1.8	318	13	Q9PTX7	Q9PTX7 lechenetron
88	8	1.8	318	17	Q27264	Q27264 mechanobact
89	8	1.8	321	12	Q91L82	Q91L82 white spot

90	8	1.8	321	12	Q8VAL2	Q8VAL2 white spot
91	8	1.8	321	12	Q913E5	Q913E5 white spot
92	8	1.8	325	10	Q8S6E7	Q8S6E7 oryza sativ
93	8	1.8	326	10	Q8S672	Q8S672 oryza sativ
94	8	1.8	333	5	Q8LJW7	Q8LJW7 sorghum bic
95	8	1.8	335	10	Q8TIF9	Q8TIF9 dicystostell
96	8	1.8	339	4	Q8NC04	Q8NC04 homo sapien
97	8	1.8	340	4	Q8IXQ4	Q8IXQ4 homo sapien
98	8	1.8	346	11	Q8K2V8	Q8K2V8 mus musculu
99	8	1.8	351	4	Q9C0P9	Q9C0P9 homo sapien
100	8	1.8	354	4	Q9BXC3	Q9BXC3 homo sapien

ALIGNMENTS

RESULT 1

Q9HCC4 PRELIMINARY; PRT; 455 AA.

AC Q9HCC4; 01-MAR-2001 (TEMBLrel. 16, Created)

DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)

DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)

DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper containing kinase AZK) (Mixed lineage kinase)-related kinase MRK-beta).

GN MLTK.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Geron I.; Adachi W.; Nishida E.;

RT "Identification and Characterization of a Novel MAP Kinase Kinase, MLTK.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RA Strausberg R.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Acton S.;

RT "MLK-mixed lineage kinase.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA MEDLINE=21950776; PubMed=11836244;

RT "Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;

RT "WRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in gamma-Radiation-induced Cell Cycle Arrest.";

RL J. Biol. Chem. 277:13873-13882(2002).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

EMBL: AB049734; BAB16445.1; -

EMBL: BC001401; AAH01401.1; -

EMBL: AF25454; AAK1615.1; -

EMBL: AF480462; AAL85892.1; -

HSSP: P12931; 1FMK.

DR InterPro: IPR000719; Prot_kinase.

DR InterPro: IPR002290; Ser_Thr_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; pkinase; 1.

DR PRINTS: PR00109; TYRKINASE.

DR PRODOM: PD000001; Prot_kinase; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 455 AA; 51582 MW; E87DB84AD58B752 CRC64;

Query Match 100.0%; Score 455; DB 4; Length 455;

Best local Similarity 100.0%; Pred. No. 0;

Matches	455;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MSLSGASFVQIKFDDIQFENC	GGSGFGSVYAAKNI	ISQKEVA	VYKKLL	IKIEKAE	LTSLV	60	
Db	1	MSLSGASFVQIKFDDIQFENC	GGSGFGSVYAAKNI	ISQKEVA	VYKKLL	IKIEKAE	LTSLV	60	
QY	61	SHRNITQFYGVILEPPNYIV	TEYASL	GLYVYINS	NRSEEDMD	IMTATD	VAKGMHY	120	
Db	61	SHRNITQFYGVILEPPNYIV	TEYASL	GLYVYINS	NRSEEDMD	IMTATD	VAKGMHY	120	
QY	121	LHMEAPVKVIRHDLKSRNV	IADGV	LKICDFGAS	FHNHTTMS	IVGTF	PMNAPEV	180	
Db	121	LHMEAPVKVIRHDLKSRNV	IADGV	LKICDFGAS	FHNHTTMS	IVGTF	PMNAPEV	180	
QY	181	LPVSECTDTYSYGVVLEML	TRVPEK	GLGEGV	QVAMLV	EKERL	LTIPSS	CRRSFELH	240
Db	181	LPVSECTDTYSYGVVLEML	TRVPEK	GLGEGV	QVAMLV	EKERL	LTIPSS	CRRSFELH	240
QY	241	QCWEADAKRPSFKQIIS	ILESMS	NDTSLP	DKNSFL	HNKAE	MRCIEAT	LERLKLERD	300
Db	241	QCWEADAKRPSFKQIIS	ILESMS	NDTSLP	DKNSFL	HNKAE	MRCIEAT	LERLKLERD	300
QY	301	LSFKQELKERRRRLKMW	QKLT	EQSNTPL	LLPLA	RMSEESY	FBSKTE	SSNAEMSCQI	360
Db	301	LSFKQELKERRRRLKMW	QKLT	EQSNTPL	LLPLA	RMSEESY	FBSKTE	SSNAEMSCQI	360
QY	361	TATSNCEGHGMP	PSLOAM	LMGFGDI	FSNKA	GAVHSG	MQIN	MAKONS	KTTSGRRBK
Db	361	TATSNCEGHGMP	PSLOAM	LMGFGDI	FSNKA	GAVHSG	MQIN	MAKONS	KTTSGRRBK
QY	421	KVNMAIGFSDP	FLSEGD	DDDDDD	GEEED	NDMDNSE	455		
Db	421	KVNMAIGFSDP	FLSEGD	DDDDDD	GEEED	NDMDNSE	455		

RESULT 2

Q9NVE9 PRELIMINARY; PRT; 800 AA.

AC Q9NVE9; 01-OCT-2000 (TEMBLrel. 15, Created)

DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)

DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed lineage kinase-related kinase MRK-alpha).

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA McNea U.J.; Frima N.; Diamond T.E.; Dower S.K.; Guesdon F.;

RT "Cloning and Characterization of AZK, a mixed lineage kinase containing a sterile-alpha motif.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=21950776; PubMed=11836244;

RT "Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;

RT "WRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in gamma-Radiation-induced Cell Cycle Arrest.";

RL J. Biol. Chem. 277:13873-13882(2002).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

EMBL: AF251441; AAF65822.1; -

EMBL: AF480461; AAL85891.1; -

HSSP: P12931; 1FMK.

DR InterPro: IPR000719; Prot_kinase.

DR InterPro: IPR001660; SAM.

DR InterPro: IPR002290; Ser_Thr_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; pkinase; 1.

DR PRINTS: PR00109; TYRKINASE.

DR PRODOM: PD000001; Prot_kinase; 1.

DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91181 MW; B289D836EC52E295 CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKWIISQKEVAVKKLKIEKEAETLSVL 60
 DB 1 MSSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKWIISQKEVAVKKLKIEKEAETLSVL 60
 QY 61 SHRNITIQFVGVILLEPPNYGIVTEYASLGSLYDINSRSEEDMDHIMTATDVAKGMHY 120
 DB 61 SHRNITIQFVGVILLEPPNYGIVTEYASLGSLYDINSRSEEDMDHIMTATDVAKGMHY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIADGVLTICDPGASRFHNHTTMSLVGTFPMAPAEVLOS 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIADGVLTICDPGASRFHNHTTMSLVGTFPMAPAEVLOS 180
 QY 181 LPVSETCDITYSGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPSPFAELH 240
 DB 181 LPVSETCDITYSGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPSPFAELH 240
 QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLTERD 300
 DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLTERD 300
 QY 301 LSFKEQLKERERRLKMWEOKLTQOSNTPL 331
 DB 301 LSFKEQLKERERRLKMWEOKLTQOSNTPL 331

RESULT 3

Q9NYL2

ID Q9NYL2 PRELIMINARY; PRT; 800 AA.

AC Q9NYL2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Mixed lineage kinase.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20384179; PubMed=10924358;
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,
 RA Chou C.K., Yang J.J.;
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein
 RT containing a leucine-zipper and a sterile-alpha motif."
 RT Biochem. Biophys. Res. Commun. 274:811-816(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF238255; AAF63490.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00454; SAM; 1_kinase; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKWIISQKEVAVKKLKIEKEAETLSVL 60
 DB 1 MSSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKWIISQKEVAVKKLKIEKEAETLSVL 60
 QY 61 SHRNITIQFVGVILLEPPNYGIVTEYASLGSLYDINSRSEEDMDHIMTATDVAKGMHY 120
 DB 61 SHRNITIQFVGVILLEPPNYGIVTEYASLGSLYDINSRSEEDMDHIMTATDVAKGMHY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIADGVLTICDPGASRFHNHTTMSLVGTFPMAPAEVLOS 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIADGVLTICDPGASRFHNHTTMSLVGTFPMAPAEVLOS 180
 QY 181 LPVSETCDITYSGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPSPFAELH 240
 DB 181 LPVSETCDITYSGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPSPFAELH 240
 QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLTERD 300
 DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLTERD 300
 QY 301 LSFKEQLKERERRLKMWEOKLTQOSNTPL 331
 DB 301 LSFKEQLKERERRLKMWEOKLTQOSNTPL 331

RESULT 4

Q9HCC5

ID Q9HCC5 PRELIMINARY; PRT; 800 AA.

AC Q9HCC5;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE MLTK-alpha.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21264927; PubMed=11042189;
 RA Gotch I., Adachi M., Nishida E.;
 RT "Identification and Characterization of a Novel MAP Kinase Kinase
 RT Kinase, MLTK."
 RT J. Biol. Chem. 276:4276-4286(2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AB049733; BAB16444.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00454; SAM; 1_kinase; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKWIISQKEVAVKKLKIEKEAETLSVL 60

```

Db      1 MSLSGASFWOIKRDDLOPFENCGGSGFSGVYRAKWSIQDKEVAVKKLKTEKEAEILLSVL 60
Qy      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120
Db      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120
Qy      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Qy      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Db      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Qy      241 QCWEADAKRRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATELRLKLE 300
Db      241 QCWEADAKRRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATELRLKLE 300
Qy      301 LSFKEQLKERERRLKMEQKLTQESNTPL 331
Db      301 LSFKEQLKERERRLKMEQKLTQESNTPL 331

```

RESULT 5

```

Q9HDD2 ID Q9HDD2 PRELIMINARY; PRT; 800 AA.
AC Q9HDD2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Placuble mixed-lineage kinase protein.
GN MLTKLK.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid organ;
RA Abe Y., Ueda N.;
RT "Placuble mixed-lineage kinase derived from LAK cell.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB030034; BABI2040.1; -.
DR HSSP: P12931; 1FMK.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50105; SAM_DOMAIN; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 9155 MW; B2814509EC54B07A CRC64;

```

Query Match 72.7%; Score 331; DB 4; Length 800;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MSLSGASFWOIKRDDLOPFENCGGSGFSGVYRAKWSIQDKEVAVKKLKTEKEAEILLSVL 60
Db      1 MSLSGASFWOIKRDDLOPFENCGGSGFSGVYRAKWSIQDKEVAVKKLKTEKEAEILLSVL 60
Qy      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120
Db      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120

```

```

Qy      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Qy      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Db      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Qy      241 QCWEADAKRRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATELRLKLE 300
Db      241 QCWEADAKRRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATELRLKLE 300
Qy      301 LSFKEQLKERERRLKMEQKLTQESNTPL 331
Db      301 LSFKEQLKERERRLKMEQKLTQESNTPL 331

```

RESULT 6

```

Q9ESL3 ID Q9ESL3 PRELIMINARY; PRT; 454 AA.
AC Q9ESL3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21264927; PubMed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB049732; BABI6443.1; -.
DR HSSP: P12931; 1FMK.
DR MGD: MGI:1931274; Zak.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

```

Query Match 36.5%; Score 166; DB 11; Length 454;
 Best Local Similarity 99.6%; Pred. No. 3; 5e-163;
 Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MSLSGASFWOIKRDDLOPFENCGGSGFSGVYRAKWSIQDKEVAVKKLKTEKEAEILLSVL 60
Db      1 MSLSGASFWOIKRDDLOPFENCGGSGFSGVYRAKWSIQDKEVAVKKLKTEKEAEILLSVL 60
Qy      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120
Db      61 SHRNIIQFYGVILLEPPNYGIVTEYASIGSLYDYINSNRSEEMDMHIMTATDVAKGMHY 120
Qy      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Db      121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFFPMMAPEVIO 180
Qy      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Db      181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSPFAELH 240
Qy      241 QCWEADAKRRPSFKQIISILESMSNDT 267

```

Db 241 QCWEADAKKRPSPKQIISLESMSNDT 267

|||||

RESULT 7
Q9ESL4 PRELIMINARY; PRT; 802 AA.
ID Q9ESL4
AC Q9ESL4
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Sterile-alpha motif and leucine zipper containing kinase
DE AZK.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; PubMed=11042189;
RA Gotch I., Adachi M., Nishida E.,
RT "Identification and Characterization of a Novel MAP Kinase Kinase, MLTK."
RL J. Biol. Chem. 276:4276-4286(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049731; BAB16442.1; -
DR EMBL; BC023718; AAH23718.1; -
DR HSSP; P12931; IFMK.
DR MGD; MGJ:1931274; Zak.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001660; SAM_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; D43IDF8312A43CC CRC64;

Query Match 36.5%; Score 166; DB 11; Length 802;
Best Local Similarity 99.6%; Pred. No. 5.9e-163;
Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSLSGSPFOIKRPDDIOFPENCGGSGFSGSVYRAKMTISQDEKVAVKKLTKIEKAELISVL 60
DB 1 MSSLSGSPFOIKRPDDIOFPENCGGSGFSGSVYRAKMTISQDEKVAVKKLTKIEKAELISVL 60
QY 61 SHRNIIQFYGVILPEPNYGVITEYASLSGLYDYNINRSREEMDMIMTATVAKGMHY 120
DB 61 SHRNIIQFYGVILPEPNYGVITEYASLSGLYDYNINRSREEMDMIMTATVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVKICDPGASRPHNHTTMSLVGTFFPMAPEVIOG 180
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVKICDPGASRPHNHTTMSLVGTFFPMAPEVIOG 180
QY 181 LPPVSECDITYSYGVVLMEMLTREVPPKGLGQVAMLYVEKNRSLTPSSCPSPFAELH 240
DB 181 LPPVSECDITYSYGVVLMEMLTREVPPKGLGQVAMLYVEKNRSLTPSSCPSPFAELH 240
QY 241 QCWEADAKKRPSPKQIISLESMSNDT 267
DB 241 QCWEADAKKRPSPKQIISLESMSNDT 267

RESULT 8

Q9BR73 PRELIMINARY; PRT; 289 AA.
ID Q9BR73
AC Q9BR73
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Sterile-alpha motif and leucine zipper containing kinase AZK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=2354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK045444; BAC32371.1; -
SQ SEQUENCE 289 AA; 32872 MW; 550E5651066A0463 CRC64;

Query Match 22.9%; Score 104; DB 11; Length 289;
Best Local Similarity 100.0%; Pred. No. 5.2e-99;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLSGSPFOIKRPDDIOFPENCGGSGFSGSVYRAKMTISQDEKVAVKKLTKIEKAELISVL 60
DB 1 MSSLSGSPFOIKRPDDIOFPENCGGSGFSGSVYRAKMTISQDEKVAVKKLTKIEKAELISVL 60
QY 61 SHRNIIQFYGVILPEPNYGVITEYASLSGLYDYNINRSREEMDM 104
DB 61 SHRNIIQFYGVILPEPNYGVITEYASLSGLYDYNINRSREEMDM 104

RESULT 9
Q90ZY8 PRELIMINARY; PRT; 371 AA.
ID Q90ZY8
AC Q90ZY8
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Protein kinase Npk.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, ZNPK, from the zebrafish."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF265343; AAK52416.1; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8A8B20D296 CRC64;

Query Match 10.8%; Score 49; DB 13; Length 371;
Best Local Similarity 100.0%; Pred. No. 7.7e-42;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 HTTHMSLVGTFFPMAPEVIOQLPVSRTCDITYSGVVLWMLTREVPPKGLG 208

DB 191 HTTPSLVGFPMAPVIOSLPVSERCDYISGVLMENLTREVPFKG 239

```

RESULT 10
Q8BUD3 PRELIMINARY; PRT; 109 AA.
AC Q8BUD3;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MLTK-beta (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK085766; BAC39534.1; -
FT NON_TER
SQ SEQUENCE 109 AA; 11627 MW; A3742E28A48E3E39 CRC64;

Query Match 5.5%; Score 25; DB 11; Length 109;
Best Local Similarity 100.0%; Pred. No. 2,1e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 KRRGKVMALGFSDPDLSEGD 440
DB 71 KRRGKVMALGFSDPDLSEGD 95

RESULT 11
Q001700 PRELIMINARY; PRT; 855 AA.
AC Q001700;
DT 01-JUN-1997 (TREMBlrel. 04, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 95.1 kDa protein F33E2.2.
GN F33E2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL022593; CA018635.2; -
DR EMBL; Z84574; CAB06544.2; JOINED.
DR EMBL; AL022593; CAB06544.2; JOINED.
DR Wormped; F33E2.2; CE23702.
DR InterPro; IPR000719; Prot. kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 855 AA; 95786 MW; F44DD253CB7D95A CRC64;

```

```

Query Match 2.4%; Score 11; DB 5; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GVLVEMLTRE 203
DB 235 GVLVEMLTRE 245

RESULT 12
Q9VW24 PRELIMINARY; PRT; 977 AA.
AC Q9VW24;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG8789 protein (LD14856p).
GN CG8789.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foelel A., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svaykaskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

```

RA Celinker S.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AA003516; AAF49129.1; -;
 DR EMBL: AY094787; AAM1140.1; -;
 DR HSSP: P08631; IAD5.
 DR FlyBase: FBgn0036896; CG8789.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_Thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase.1.
 DR PROSITE: PSS0011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR ATP-binding_kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 977 AA; 109567 MW; 5CD235939C14BD4E CRC64;

Query Match 2.4%; Score 11; DB 5; Length 977;
 Best Local Similarity 100.0%; Pred. No. 0.05; Mismatches 0; Gaps 0;
 Matches 11; Conservative 0; Indels 0;

QY 191 SYGVLMEMLT 201
 DB 331 SYGVLMEMLT 341

RESULT 13
 ID Q9SMA2 PRELIMINARY; PRT; 141 AA.
 AC Q9SMA2;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DE 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE Zwh0007.1.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;
 RX SEQUENCE FROM N.A.
 RP STRAIN=cv. indica;
 RA Hong G., Zhao W.;
 RT "Oryza sativa genomic DNA, chromosome 4, clone: b6015."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL117264; CAB55391.1; -;
 DR Gramene: Q9SMA2; -;
 SQ SEQUENCE 141 AA; 15696 MW; 1C608AB0459DD77D CRC64;

Query Match 2.0%; Score 9; DB 10; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
 DB 26 DDDDDGEE 34

RESULT 14
 ID Q8ST12 PRELIMINARY; PRT; 205 AA.
 AC Q8ST12;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Hypothetical protein 6/101
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 NCBI_TaxID=44689;
 RX SEQUENCE FROM N.A.
 RP STRAIN=AK4;

RA Gloeckner G., Eichinger L., Szafranski K., Pachbat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC115578; AAL2211.1; -;
 DR InterPro: IPR002833; UPF0099.
 DR Pfam: PF01981; UPF0099.1.
 DR ProDom: PD010667; UPF0099.1.
 DR TIGRFAMs: TIGR00283; TIGR00283.1.
 KW Hypothetical protein.
 SQ SEQUENCE 205 AA; 22960 MW; 7E3D24AD2290C4D0 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 205;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDD 443
 DB 56 EGDGDDDD 64

RESULT 15
 ID O15318 PRELIMINARY; PRT; 233 AA.
 AC O15318;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DE 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE RNA polymerase III subunit.
 GN RRC32.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=97315201; PubMed=9171375;
 RA Wang Z., Roeder R.G.;
 RT "Three human RNA polymerase III-specific subunits form a subcomplex
 RT with a selective function in specific transcription initiation."
 RL Gene Dev. 11:1315-1326(1997).
 DR EMBL: U93868; AAB53676.1; -;
 SQ SEQUENCE 233 AA; 27299 MW; 16EAB6A5F5002FF CRC64;

Query Match 2.0%; Score 9; DB 4; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDD 443
 DB 173 EGDGDDDD 181

RESULT 16
 ID Q98182 PRELIMINARY; PRT; 297 AA.
 AC Q98182;
 DT 01-FEB-1997 (T-EMBLrel. 02, Created)
 DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE MC011L.
 GN MC011L.
 OS Moluscum contagiosum virus subtype 1 (MCV1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Moluscipoxvirus.
 NCBI_TaxID=10280;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=96325459; PubMed=8670425;
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
 RA Moss B.;
 RT "Genome sequence of a human tumorigenic poxvirus: Prediction of

RT specific host response-evasion genes."

RL Science 273:813-816(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Senkevich T.G., Bugert J.J., Sialer J.R., Koonin E.V., Darai G.,

RA Moss B.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U60315; AAC55139.1; -;

SO SEQUENCE 297 AA; 31500 MW; 86199936CD8A781 CRC64;

Query Match 2.0%; Score 9; DB 12; Length 297;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDDG 444

DB 206 GDDDDDDDDG 214

RESULT 17

08GRY7 PRELIMINARY; PRT; 321 AA.

AC 08GRY7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE bZIP with a Ring-finger motif.

GN ASTRAY.

OS Lotus japonicus.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosoid I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

OC NCBI_TaxID=34305;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Gifu, and Miyakojima;

RC MEDLINE=2317398; PubMed=12397181;

RA Nishimura R., Ohmori M., Fujita H., Kawaguchi M.;

RT "A Lotus basic leucine zipper protein with a RING-finger motif

negatively regulates the developmental program of nodulation."

RL Proc. Natl. Acad. Sci. U.S.A. 99:15206-15210(2002).

DR EMBL: AB092677; BAC20318.1; -;

DR EMBL: AB092678; BAC20319.1; -;

DR EMBL: AB092679; BAC20320.1; -;

SO SEQUENCE 321 AA; 35556 MW; 9F64AFLA4D620262 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 321;

Best Local Similarity 100.0%; Pred. No. 2.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443

DB 98 EGDGDDDDDD 106

RESULT 18

039585 PRELIMINARY; PRT; 408 AA.

AC 039585;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Protein kinase.

OS Chlamydomonas reinhardtii.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonas.

OC NCBI_TaxID=3055;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=21gr(+);

RC MEDLINE=96133873; PubMed=8552645;

RA Kurvari V., Zhang Y., Luo Y., Snell W.J.;

RT "Molecular cloning of a protein kinase whose phosphorylation is

RT regulated by genetic adhesion during Chlamydomonas fertilization."

RL Proc. Natl. Acad. Sci. U.S.A. 93:39-43(1996).

DR EMBL: U36196; AAA69556.1; -;

DR InterPro: IPR000644; CBS_domain.

DR Pfam: PF00571; CBS; 3.

DR SMART: SM00116; CBS; 3.

KW kinase.

SO SEQUENCE 408 AA; 44946 MW; DBDE6B30C49CF2 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 408;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443

DB 393 EGDGDDDDDD 401

RESULT 19

08S5D1 PRELIMINARY; PRT; 453 AA.

AC 08S5D1;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypothetical 49.5 kDa protein.

GN OUI341F06.5.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OC NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RC McCombie W.R., Spiegel L., de la Bastide M., Preston R., Ferraro K.,

RA Kuit K., Nascimento L., Zutavern T., Ballja V., Bell M., Baker J.,

RA Dike S., Miller B., Katzenberger F., Miller S., King L., Yang C.,

RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone

OUI341F06, from chromosome 10, complete sequence."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC116926; AAM08885.1; -;

DR Gramene: 08S5D1; -;

KW Hypothetical protein.

SO SEQUENCE 453 AA; 49540 MW; 3661B92E92399644 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 453;

Best Local Similarity 100.0%; Pred. No. 3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDDG 444

DB 277 GDDDDDDDDG 285

RESULT 20

08S5G9 PRELIMINARY; PRT; 460 AA.

AC 08S5G9;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypothetical 50.4 kDa protein.

GN OSJNBA0091J06.10.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OC NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA McCombie W.R., de la Baetide M., Spiegel L., Preston R., Ferraro K.,
 RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
 RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.,
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
 OSJNBa009106, from chromosome 10, complete sequence."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC113338; AAM08654.1; -.
 DR Gramene; Q8S563; -.
 KW Hypothetical protein.
 SQ SEQUENCE 460 AA; 50423 MW; 4624B7D90845941A CRC64;

Query Match 2.0%; Score 9; DB 10; Length 460;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDG 444
 DB 284 GDDDDDDG 292

RESULT 21

ID Q91YV1 PRELIMINARY; PRT; 502 AA.
 AC Q91YV1;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Similar to activin A receptor, type II-like 1.
 GN ACVRL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: BC014291; AAH14291.1; -.
 DR MGD; MGI:1338946; Acvrl1.
 DR InterPro; IPR000472; Activin rec.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002280; Ser_thr_kinase.
 DR InterPro; IPR003605; TGFbeta_GS.
 DR Pfam; PF01064; Activin_recpt. 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
 KW Transferase.
 SQ SEQUENCE 502 AA; 56483 MW; 800E8254703AE875 CRC64;

Query Match 2.0%; Score 9; DB 11; Length 502;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
 DB 327 HRDLKSRNV 335

RESULT 22

ID Q91YR0 PRELIMINARY; PRT; 502 AA.
 AC Q91YR0;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Activin A receptor, type II-like 1.

GN ACVRL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: BC015083; AAH15083.1; -.
 DR MGD; MGI:1338946; Acvrl1.
 DR InterPro; IPR000472; Activin rec.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002280; Ser_thr_kinase.
 DR InterPro; IPR003605; TGFbeta_GS.
 DR Pfam; PF01064; Activin_recpt. 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
 KW Transferase.
 SQ SEQUENCE 502 AA; 56519 MW; 439510D3CC740D65 CRC64;

Query Match 2.0%; Score 9; DB 11; Length 502;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
 DB 327 HRDLKSRNV 335

RESULT 23

ID Q63559 PRELIMINARY; PRT; 504 AA.
 AC Q63559;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Serine/threonine kinase receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;
 RX MEDLINE=96198306; PubMed=8928814;
 RA Panchenko M.P., Williams M.C., Brody J.S., Yu Q.;
 RT "Type I receptor serine-threonine kinase preferentially expressed in
 RT pulmonary blood vessels.";
 RT Am. J. Physiol. 270:L547-L558 (1996).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: I36088; AAC37705.1; -.
 DR InterPro; IPR000472; Activin rec.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002280; Ser_thr_kinase.
 DR InterPro; IPR003605; TGFbeta_GS.
 DR Pfam; PF01064; Activin_recpt. 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 504 AA; 56703 MW; B40EA3075223C8F CRC64;

Query Match 2.0%; Score 9; DB 11; Length 504;

Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
|||||

Db 329 HRDLKSRNV 337

RESULT 24

Q8MYF3 PRELIMINARY; PRT; 513 AA.

AC Q8MYF3; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Presalk protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117079; AM43731.1; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KM ATP-binding; Transferase.

QY SEQUENCE 513 AA; 57705 MW; CF97F7E583A96107 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 513;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSVYRA 33
|||||

Db 51 GSFGSVYRA 59

RESULT 25

Q9FG45 PRELIMINARY; PRT; 607 AA.

AC Q9FG45; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Smiliarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Kaneo T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP00607; BAB10964.1; -
QY SEQUENCE 607 AA; 70548 MW; 6A93AF08FF94634 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 607;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDG 444
|||||

Db 586 GDDDDDDG 594

RESULT 26

Q9LKB4 PRELIMINARY; PRT; 611 AA.

AC Q9LKB4; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Genomic DNA, chromosome 3, TAC clone:K15M2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II."
RT Sequence features of the regions of 4,251,655 bp covered by ninety P1,
TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
DR EMBL; AP000370; BAA97053.1; -
QY SEQUENCE 611 AA; 71221 MW; ACE12328A9429778 CRC64;

QY 437 DDDDDDDG 445
|||||

Db 446 DDDDDDDG 454

Query Match 2.0%; Score 9; DB 10; Length 611;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 27

Q9GRH2 PRELIMINARY; PRT; 616 AA.

AC Q9GRH2; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Tyrosine kinase (EC 2.7.1.112) (Tyrosine-protein kinase receptor).
OS Sycon raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcarenea; Leucosoleniida;
OC Sycettidae.
NCBI_TaxID=56443;
RN [1]
RP SEQUENCE FROM N.A.
RC Skotokhod A., Gamulin V., Gundacker D., Kavean V., Mueller I.M.,
Mueller W.E.G.;
RT "Origin of insulin receptor tyrosine kinases in marine sponges";
RL Biol. Bull. 197:198-206(2000).
CC -I- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
DR EMBL; Y17877; CAC14729.1; -

DR HSSP; P06213; 11RK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002011; RTKinaseI.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00219; TyrcK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR ATP-binding; Glycoprotein_kinase; Phosphorylation; Receptor;
 DR Transferase; Transmembrane; Tyrosine-protein kinase
 KW SEQUENCE 616 AA; 69477 MW; D573241383DBE1E4 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 616;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVNMEM 199
 DB 425 SYGVNMEM 433

RESULT 28
 ID Q9A0J5 PRELIMINARY; PRT; 652 AA.
 AC Q9A0J5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA_ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase
 [NAD+]).
 GN LIG OR SPY0751.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
 Primeaux C., Szatec S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P.,
 Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RL -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER
 LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
 STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
 THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
 DAMAGED DNA (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NM(+) = AMP + NICOTINAMIDE NUCLEOTIDE +
 (DEOXYRIBONUCLEOTIDE) (N+M).
 CC (DEOXYRIBONUCLEOTIDE) (N+M).
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
 CC EMBL; AEO06527; AA033695.1; -.
 DR HSSP; O87703; 1B04.
 DR InterPro; IPR001357; BRCT.
 DR InterPro; IPR001679; DNAligase.
 DR InterPro; IPR004150; DNAligase_OB.
 DR InterPro; IPR000445; Hnh.
 DR InterPro; IPR003583; Hnh_1.
 DR InterPro; IPR004149; Znf_DNAligase_C4.
 DR Pfam; PF00533; BRCT; 1.
 DR Pfam; PF01653; DNA_ligase_N; 1.
 DR Pfam; PF03120; DNA_ligase_OB; 1.
 DR Pfam; PF03119; DNA_ligase_ZBD; 1.
 DR ProDom; PD003944; DNAligase; 1.
 DR SMART; SM00292; BRCT; 1.
 DR SMART; SM00278; Hnh1; 3.

DR SMART; SM00532; LIGANC; 1.
 DR TIGRPAWS; TIGR00575; dntj; 1.
 DR PROSITE; PS01722; BRCT; 1.
 DR PROSITE; PS01055; DNA_LIGASE_N; 1.
 DR PROSITE; PS01056; DNA_LIGASE_N2; 1.
 DR DNA repair; DNA replication; ligase; NAD; Complete proteome.
 KW SEQUENCE 652 AA; 72420 MW; B88D93C9FP638BE5 CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59
 DB 307 EKEAEILSV 315

RESULT 29
 ID Q8P1L1 PRELIMINARY; PRT; 652 AA.
 AC Q8P1L1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative DNA_ligase.
 GN LIG OR SPY18.0810.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=186103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Yeasty L.G., Musser J.M.;
 "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever
 outbreaks";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 RL EMBL; AEO10012; AL97474.1; -.
 DR InterPro; IPR001357; BRCT.
 DR InterPro; IPR001679; DNAligase.
 DR InterPro; IPR004150; DNAligase_OB.
 DR InterPro; IPR000445; Hnh.
 DR InterPro; IPR004149; Znf_DNAligase_C4.
 DR Pfam; PF00533; BRCT; 1.
 DR Pfam; PF01653; DNA_ligase_N; 1.
 DR Pfam; PF03120; DNA_ligase_OB; 1.
 DR Pfam; PF03119; DNA_ligase_ZBD; 1.
 DR ProDom; PD003944; DNAligase; 1.
 DR TIGRPAWS; TIGR00575; dntj; 1.
 DR PROSITE; PS01722; BRCT; 1.
 DR PROSITE; PS01055; DNA_LIGASE_N; 1.
 DR PROSITE; PS01056; DNA_LIGASE_N2; 1.
 DR ligase; Complete proteome; Hypothetical protein.
 KW SEQUENCE 652 AA; 72394 MW; E89873C9FP66444E CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59
 DB 307 EKEAEILSV 315

RESULT 30
 ID Q8K831 PRELIMINARY; PRT; 652 AA.
 AC Q8K831;

```

DT 01-OCT-2002 (T-EMBLrel.. 22, Created)
DT 01-OCT-2002 (T-EMBLrel.. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel.. 23, Last annotation update)
DE Putative DNA ligase.
GN LIG OR SPYM3_0491.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=2133808; PubMed=12122206;
RA Bares S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlivert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RL EMBL; AE014146; AM979098.1; -.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001679; DNALigase.
DR InterPro; IPR004150; DNA_ligase_OB.
DR InterPro; IPR000445; HhH.
DR InterPro; IPR003583; HHH.1.
DR InterPro; IPR004149; Znf_DNALigase_C4.
DR Pfam; PF00553; BRCT.1.
DR Pfam; PF01653; DNA_ligase_N; 1.
DR Pfam; PF03120; DNA_ligase_OB; 1.
DR Pfam; PF03119; DNA_ligase_ZB; 1.
DR ProDom; PD003944; DNALigase; 1.
DR SMART; SM00292; BRCT; 1.
DR SMART; SM00278; HhH1; 3.
DR TIGRFAMs; TIGR00575; dnlj; 1.
DR PROSITE; PS01072; BRCT; 1.
DR PROSITE; PS01055; DNA_LIGASE_N1; 1.
DR PROSITE; PS01056; DNA_LIGASE_N2; 1.
DR Ligase; Complete proteome; Hypothetical protein.
KW SEQUENCE 652 AA; 72392 MW; C3617947F8978B4 CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59
Db 307 EKEAEILSV 315

RESULT 31
Q8ESW1 PRELIMINARY; PRT; 652 AA.
ID Q8ESW1;
AC Q8ESW1;
DT 01-MAR-2003 (T-EMBLrel.. 23, Created)
DT 01-MAR-2003 (T-EMBLrel.. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel.. 23, Last annotation update)
DE Hypothetical protein.
DE Hypothetical protein.
GN GBS0868.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaeser T., Rusanok C., Buchrieser C., Chevallier F., Frangoul L.,
RA Maadock T., Zouine M., Couve E., Lailouli L., Poyart C., Trieru-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing

```

```

RT Invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766847; CAD46512.1; -.
DR Sagaliet; gbs0868; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 652 AA; 72683 MW; 1B998628DB17734 CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59
Db 307 EKEAEILSV 315

RESULT 32
Q8E084 PRELIMINARY; PRT; 652 AA.
ID Q8E084;
AC Q8E084;
DT 01-MAR-2003 (T-EMBLrel.. 23, Created)
DT 01-MAR-2003 (T-EMBLrel.. 23, Last sequence update)
DE DNA ligase, NAD-dependent.
GN LIGA OR SAG0850.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tetteiln H., Masiyanti V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarcelll M., Mora M.,
RA Iacobini E.T., Bretoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Frazer C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae."
RT Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
RL EMBL; AE014230; AM99737.1; -.
DR TIGR; SAG0850; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 652 AA; 72579 MW; 6CEB70676C2C406F CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59
Db 307 EKEAEILSV 315

RESULT 33
Q12500 PRELIMINARY; PRT; 764 AA.
ID Q12500;
AC Q12500;
DT 01-NOV-1996 (T-EMBLrel.. 01, Created)
DT 01-NOV-1996 (T-EMBLrel.. 01, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel.. 23, Last annotation update)
DE Chromosome XII reading frame ORF YLR114C.
GN ERF4 OR L9354.8 OR L2941 OR YLR114C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA Vernasselt P., Voet M., Volckaert G.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Geisel C.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Ravello A., Fulton L., Galling S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
 RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Meneses S.,
 RA Miller N., Nhan M., Paulley A., Peluso D., Ricken L., Riles L.,
 RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Waterston R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U73286; CAA97681.1; -;
 DR EMBL; U53878; AAB67559.1; -;
 DR EMBL; X89514; CAA61692.1; -;
 DR SGD; S0004104; YLR114C.
 SQ SEQUENCE 764 AA; 86425 MW; 159A00CF933A5C4A CRC64;

Query Match 2.0%; Score 9; DB 3; Length 764;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445
 |||||
 DB 628 DDDDDDDGE 636

RESULT 34
 Q9XBP8 PRELIMINARY; PRT; 789 AA.
 AC Q9XBP8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Serine/threonine kinase PKN11.
 GN PKN11.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Delta proteobacteria; Myxococcales;
 OC Cytochromatellaceae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DZf1;
 AC Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,
 RA Munoz-Dorado J., Fares-Vidal E., Inouye M.;
 RL "Sequence Analysis of 13 Eukaryotic-like Protein Ser/Thr Kinases of
 RT Myxococcus xanthus, a Developmental Bacterium and Significance of
 RT Their Coexistence with Protein His Kinases."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF159690; AAD42854.1; -;
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.

SQ SEQUENCE 789 AA; 81613 MW; EE1E764F0BD77C63 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 789;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GVLWEMLT 201
 |||||
 DB 210 GVLWEMLT 218

RESULT 35
 Q8K062 PRELIMINARY; PRT; 840 AA.
 AC Q8K062;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to RIKEN CDNA 4921513020 gene (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034074; AAH34074.1; -;
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000408; Reg_Chr_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PRO0109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TYRKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00626; RCC1_2; 1.
 KW ATP-binding; Transferase.
 FT NON_TER

SQ SEQUENCE 840 AA; 94784 MW; A3AFDA4D5F712109 CRC64;

Query Match 2.0%; Score 9; DB 11; Length 840;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSVYRA 33
 |||||
 DB 201 GSFGSVYRA 209

RESULT 36
 Q9ZSM8 PRELIMINARY; PRT; 856 AA.
 AC Q9ZSM8;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE E2A1 (Atg92020/t10M13_3).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bildeau P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;
 RT "E2A1, a novel polycomb group gene from Arabidopsis thaliana."
 RL Plant Physiol. 0:0-0(1999).

```

RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bower L., Carinict P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlín-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
RA Bower L., Carinict P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlín-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR EMBL; AP100163; AAD09108.1; -.
DR EMBL; AY057477; AAL09711.1; -.
DR EMBL; AY090293; AAL90954.1; -.
DR HSSP; P10969; 1WGT.
DR InterPro; IPR001005; MYB_DNA_binding.
DR InterPro; IPR001214; SET.
DR InterPro; IPR002221; MAP.
DR Pfam; PF00856; SET; 1.
DR SMART; SM00717; SANT; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00317; 4 DISULFIDE_CORE; 1.
DR PROSITE; PS50280; SET; 1.
SQ SEQUENCE 856 AA; 95396 MW; DD4B099C936F197C CRC64;

Query Match 2.0%; Score 9; DB 10; Length 856;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
DB 17 DDDDDGEE 25

RESULT 37
ID 004246 PRELIMINARY; PRT; 898 AA.
AC 004246;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative CURVLEAF-like 1 homeotic protein (POLYCOMB group-like
DE protein).
GN T10M13.3 OR AT4G02020.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Johnson A.F., de la Bastide M., Lodhi M., Hoffman J., Hasegawa A.,
RA Gnoj L., Gottesman T., Granat S., Hameed A., Kaplan N., Schutz K.,
RA Shohdy N., Van Keulen K., Parnell L., Dedhia N., Martienssen R.,
RA McCombie W.;
RT "The sequence of the Arabidopsis thaliana T10M13 BAC."
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Prescon R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR EMBL; AF001308; AAC78694.1; -.
DR EMBL; AL161493; CAB80695.1; -.
DR HSSP; P10969; 1WGT.
DR InterPro; IPR001005; MYB_DNA_binding.
DR InterPro; IPR001214; SET.
DR InterPro; IPR002221; MAP.
DR Pfam; PF00856; SET; 1.
DR SMART; SM00717; SANT; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00317; 4 DISULFIDE_CORE; 1.
DR PROSITE; PS50280; SET; 1.
SQ SEQUENCE 898 AA; 100379 MW; 79BE740689C99C2E CRC64;

Query Match 2.0%; Score 9; DB 10; Length 898;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
DB 83 DDDDDGEE 91

RESULT 38
ID 09P8G2 PRELIMINARY; PRT; 1059 AA.
AC 09P8G2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Potassium transporter Trk1p.
GN TRK1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OC NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA Miranda M., Bash E., Slayman C.L.;
RT "Cloning and characterization of a TRK homolog from Candida albicans
RT (TRK1) by complementation in Saccharomyces."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267125; AAF72203.1; -.
DR InterPro; IPR003445; Cat transport.
DR InterPro; IPR004773; Ktrana_p_euk.
DR Pfam; PRO2386; TrkH; 1.
DR TIGRFAMs; TIGR00934; 2a38euk; 1.
SQ SEQUENCE 1059 AA; 119788 MW; 9C643485C9453747 CRC64;

Query Match 2.0%; Score 9; DB 3; Length 1059;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 445
DB 482 DDDDDGEE 490

RESULT 39
ID 09U5A8 PRELIMINARY; PRT; 1472 AA.
AC 09U5A8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Insulin receptor-like protein precursor (EC 2.7.1.112) (Tyrosine-
 DE protein kinase receptor).
 GN BIR.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OC NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary vitellogenic follicles;
 RA Lindstrom-Dinietz I., Iatrou K.;
 RT "Cloning and functional characterization of an insulin receptor-like
 RT mRNA expressed in the silkworm ovary."
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
 CC TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC EMBL: AF025542; AAF21243.1; -.
 CC HSSP: P06213; 11RK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR003961; FN_III_
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin_repeat.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002011; RTK_kinaseII.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00041; fn3; 1.
 DR Pfam: PF00057; Furin-like; 1.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Prot_kinase; 1.
 DR SMART: SM00060; FN3; 3.
 DR SMART: SM00261; FN; 1.
 DR SMART: SM00219; TYKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
 KW ATP-binding; Glycoprotein; Kinase; Phosphorylation; Receptor; Signal;
 KW Transferrase; Transmembrane; Tyrosine-protein kinase.
 FT TRANSAL 1 30 POTENTIAL.
 FT CHAIN 31 1472 INSULIN RECEPTOR-LIKE PROTEIN.
 SQ SEQUENCE 1472 AA; 164582 MW; CB6631C67906AF8E CRC64;
 QY Query Match 2.0%; Score 9; DB 5; Length 1472;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1283 SYGVLMEM 1291
 QY 191 SYGVLMEM 199
 DB 1283 SYGVLMEM 1291
 RESULT 40
 ID 081A15 PRELIMINARY; PRT; 1738 AA.
 AC 081A15;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Hypothetical protein.
 GN MAL8P1.157.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
 RA Quail M., Barrell B.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL844507; CAD51348.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 1738 AA; 205288 MW; SCBF1507654FBA4 CRC64;
 QY Query Match 2.0%; Score 9; DB 5; Length 1738;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 879 GDDDDDDG 887
 QY 436 GDDDDDDG 444
 DB 879 GDDDDDDG 887
 RESULT 41
 ID 08T5H2 PRELIMINARY; PRT; 1978 AA.
 AC 08T5H2;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Putative Tyr/Ser/Thr phosphatase.
 GN 22J3.6.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 OC NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST.
 RA Thomasova D., Ton L.O., Collins F.H., Kafatos F.C.;
 RT "Sequencing and analysis of the EMBL/GenBank/DBJ databases.
 RT chromosome 2R."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A43398; CAD28129.1; -.
 DR InterPro: IPR000387; Tyr_phosphatase.
 DR InterPro: IPR000306; Znf_FYVE.
 DR Pfam: PF01363; FYVE; 1.
 DR SMART: SM00064; FYVE; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50178; ZF_FYVE; 1.
 KW Hydrolase.
 SQ SEQUENCE 1978 AA; 211417 MW; E0D9506B19B35F5 CRC64;
 QY Query Match 2.0%; Score 9; DB 5; Length 1978;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1733 DDDDDGEE 1741
 QY 438 DDDDDGEE 446
 DB 1733 DDDDDGEE 1741
 RESULT 42
 ID 027732 PRELIMINARY; PRT; 2391 AA.
 AC 027732;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Cardamoyl phosphate synthetase II (EC 6.3.5.5) (Copper amine
 DE oxidase).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCQ27;
 RA Flores M.V.C.;
 RT "Molecular Biology of Pyrimidine Biosynthesis in the Human Malarial
 RT Parasite Plasmodium falciparum.";

RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCQ27;
 RA Flores M.V.C., O'Sullivan W.J., Stewart T.S.;
 RT "Characterization of the Cardamoyl Phosphate Synthetase Gene from
 RT Plasmodium falciparum";
 RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
 CC H(2)O(2).
 CC -1- COFACTOR: BINDS 1 COPPER ION AND 1 TOPAQUINONE PER SUBUNIT (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.
 CC EMBL; L32150; AAA29522.1; -.
 DR HSSP; P00968; 1A9X.
 DR InterPro; IPR006275; CARA_L_glu.
 DR InterPro; IPR006274; CARA_small.
 DR InterPro; IPR005483; CPase_L.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR005480; CPase_L_D3.
 DR InterPro; IPR005481; CPase_L_N.
 DR InterPro; IPR002474; CPase_sm_chain.
 DR InterPro; IPR001317; CPS_GATase.
 DR InterPro; IPR000269; CUNH_oxidase.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000991; GATase_1.
 DR InterPro; IPR004362; MGS_like.
 DR Pfam; PF00289; CPase_L_chain; 2.
 DR Pfam; PF02786; CPase_L_D2; 2.
 DR Pfam; PF02787; CPase_L_D3; 1.
 DR Pfam; PF00988; CPase_sm_chain; 1.
 DR Pfam; PF00117; GATase; 1.
 DR Pfam; PF02142; MGS; 1.
 DR PRINTS; PR00098; CPASE.
 DR PRINTS; PR00099; CPSEATASE.
 DR PRINTS; PR00096; GATASE.
 DR TIGRFAMs; TIGR01368; CPase_small; 1.
 DR TIGRFAMs; TIGR01369; CPase_1g; 1.
 DR PROSITE; PS01164; COPPER_AMINE_OXID_1; 1.
 DR PROSITE; PS00866; CPASE_1; 1.
 DR PROSITE; PS00867; CPASE_2; 2.
 DR PROSITE; PS00018; EF_HAND; 1.
 KW COPPER; Ligase; Oxidoreductase; TPO.
 SQ SEQUENCE 2391 AA; 275674 MW; 426F3C3E612FE5E2 CRC64;
 QY Query Match 2.0%; Score 9; DB 5; Length 2391;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDDGE 445
 DB 1786 DDDDDDDGE 1794

RT clone:P0434A03."
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF004299; BAC45128.1; -.
 SQ SEQUENCE 70 AA; 7558 MW; 93E47CC7E9716993 CRC64;
 QY Query Match 1.8%; Score 8; DB 10; Length 70;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 436 GDDDDDD 443
 DB 29 GDDDDDD 36

RESULT 44
 ID 015452 PRELIMINARY; PRT; 101 AA.
 AC 015452;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Protein-serine/threonine kinase (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94100173; PubMed=8274451;
 RA Schultz S.J., Nigg E.A.;
 RT "Identification of 21 novel human protein kinases, including 3 members
 RT of a family related to the cell cycle regulator nimA of Aspergillus
 RT nidulans";
 RL Cell Growth Differ. 4:821-830 (1993).
 DR EMBL; Z25430; CA80917.1; -.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.
 FT NON_TER 1
 FT NON_TER 101
 SQ SEQUENCE 101 AA; 11485 MW; E584845C18F4FCD8 CRC64;
 QY Query Match 1.8%; Score 8; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 172 WMAPEVIQ 179
 DB 92 WMAPEVIQ 99

RESULT 45
 ID 09J872 PRELIMINARY; PRT; 106 AA.
 AC 09J872;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ORF63.
 OS Spodoptera exigua nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OC NCBI_TaxID=10454;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20036646; PubMed=10567663;
 RA Ickel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RT "Sequence and organization of the spodoptera exigua multicapsid
 RT nucleopolyhedrovirus genome";
 RT J. Gen. Virol. 80:3289-3304 (1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RA IJkel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF169823; AAF33593.1; -
 SQ SEQUENCE 106 AA; 11780 MW; 599AA3E43F88023 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 106;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 DDDDDDD 443
 DB 11 DDDDDDD 18

RESULT 46
 Q8E280 PRELIMINARY; PRT; 145 AA.
 AC Q8E280;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative globin-like protein.
 GN LA3976.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 ON NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF011553; AAN51174.1; -
 KW Complete proteome.
 SQ SEQUENCE 145 AA; 16422 MW; D29B48EB7F84A8C3 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 145;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 ESKTEESN 352
 DB 136 ESKTEESN 143

RESULT 47
 Q8LIE2 PRELIMINARY; PRT; 148 AA.
 AC Q8LIE2;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE OJ1343 D04.1 protein (P0453E03.20 protein).
 GN OJ1343_D04.1 OR P0453E03.20.
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 ON NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RT clone:OJ1343 D04.1;
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC

RT clone:P0453E03.20;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003825; BAC10084.1; -
 DR EMBL; AP005452; BAC22432.1; -
 DR Gramene; Q8LIE2; -
 SQ SEQUENCE 148 AA; 15740 MW; 858BA91E2263A055 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 148;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 DB 59 DDDDDDDG 66

RESULT 48
 O55500 PRELIMINARY; PRT; 157 AA.
 AC O55500;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE UL51.
 GN UL51.
 OS Human cytomegalovirus (strain Towne).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 ON NCBI_TaxID=10363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Towne;
 RX MEDLINE=98241707; PubMed=9573236;
 RA Krosky P.M., Underwood M.R., Turk S.R., Feng K.W., Jain R.K.,
 RA Plak R.G., Westerman A.C., Biron K.K., Townsend L.B., Drach J.C.;
 RT "Resistance of human cytomegalovirus to benzimidazole ribonucleosides
 maps to two open reading frames: UL89 and UL56.";
 RL J. Virol. 72:4721-4728 (1998).
 DR EMBL; AF039234; AAC59474.1; -
 DR InterPro; IPR005208; Herpes_UL33.
 DR Pfam; PF03581; Herpes_UL33; 1.
 SQ SEQUENCE 157 AA; 16978 MW; 5999035AC484517D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 157;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 DDDDGEE 447
 DB 12 DDDDGEE 19

RESULT 49
 Q8IDW5 PRELIMINARY; PRT; 158 AA.
 AC Q8IDW5;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PPI3_0203.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B., Lennard N., Clark L., Line A., Barton A., Corton C.,
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL844509; CAD52504.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 158 AA; 18268 MW; FE788EF8723A335C CRC64;

Query Match 1.8%; Score 8; DB 5; Length 158;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 DDDDDDDG 443
 |||||
 Db 87 DDDDDDDG 94

RESULT 50

Q91902 PRELIMINARY; PRT; 161 AA.
 AC Q91902;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 GN CUN017 hypothetical protein.
 OS Culex nigripalpus baculovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.
 OX NCBI_TaxID=130556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Florida197;
 RX Atcnso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
 RA Becnel J.J., Rock D.L., Kutish G.F.;
 RT "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus.";
 RL J. Virol. 75:11157-11165(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Florida197;
 RX Atcnso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,
 RA Becnel J.J., Rock D.L., Kutish G.F.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF403738; AK94095.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 161 AA; 18553 MW; 99BA93D3FB8386DD CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
 |||||
 Db 116 DDDDDDDG 123

RESULT 51

Q9J867 PRELIMINARY; PRT; 161 AA.
 AC Q9J867;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE ORF68
 OS Spodoptera exigua nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OX NCBI_TaxID=10454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20036646; PubMed=10567663;
 RX Ickel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RT "Sequence and organization of the spodoptera exigua multicapsid
 RT nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 80:3289-3304(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ickel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF169823; AAFF33598.1; -
 DR InterPro: IPR002557; Chitin_bind_Pera.
 DR Pfam: PF01607; CBM_14; 1.
 DR SMART: SM00494; ChtBD2; 1.
 SQ SEQUENCE 161 AA; 18509 MW; AE3EA844D73A9B73 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
 |||||
 Db 142 DDDDDDDG 149

RESULT 52

Q9HWO2 PRELIMINARY; PRT; 162 AA.
 AC Q9HWO2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 GN PA4129.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Binkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber L.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Polger K.R., Kas A., Lapidis K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V., an
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004829; AAG07516.1; -
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 162 AA; 18122 MW; 094CA423A6E87795 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 162;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 PLILPLAA 336
 |||||
 Db 34 PLILPLAA 41

RESULT 53

Q9P840 PRELIMINARY; PRT; 180 AA.
 AC Q9P840;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Essential YAE1 protein.
 GN YAE1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA De Backer M.D., Logghe M., Viane J., Ioonen I., Vandoninck S.,
 RA de Hoogt R., Nelissen B., Dewaele S., Simons F., Verhaeselt P.,
 RA Contreras R., Luyten W.H.M.L.;

RT "A novel method for systematic identification of genes required for
RT growth of *Candida albicans*.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ390499; CAB77639.1; -.
SQ SEQUENCE 180 AA; 20240 MW; EC98B8B63C1676C7 CRC64;

Query Match 1.8%; Score 8; DB 3; Length 180;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 117 GDDDDDD 124

RESULT 54
ID 000193 PRELIMINARY; PRT; 183 AA.
AC 000193;
DT 01-JUL-1997 (TEMBLrel. 04, Created)

DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Small acidic protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NC NCB1_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97408516; PubMed=9263035;
RA Gong T.W., Hegeman A.D., Shin J.J., Lindberg K.H., Barald K.F.,
Lomax M.I.;

RT "Novel genes expressed in the chick oocyte during development:
RT identification using differential display of RNA.";
RL Int. J. Dev. Neurosci. 15:585-594(1997).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U51678; AAB99729.1; -.
DR EMBL; BC016352; AAH16352.1; -.
DR EMBL; BC007103; AAH07103.1; -.
DR EMBL; BC020937; AAH20937.1; -.
SQ SEQUENCE 183 AA; 20332 MW; 509BA7377B69A74 CRC64;

Query Match 1.8%; Score 8; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 118 GDDDDDD 125

RESULT 55
ID 042341 PRELIMINARY; PRT; 184 AA.
AC 042341;
DT 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Serine-threonine protein kinase (Mouse-ear cress).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
NC NCB1_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Cooke R., Landie M., Raynal M., Delseny M.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; F19996; CAA23375.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
FT 184
SQ SEQUENCE 184 AA; 20933 MW; EB93761356ACE2E0 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 137 WMAPEVIQ 144

RESULT 56
ID 090AM0 PRELIMINARY; PRT; 194 AA.
AC 090AM0;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Hypothetical 22.3 kDa protein.
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
NC NCB1_TaxID=92652;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20112891; PubMed=10644828;
RA van Hulten M.C., Tsai M.F., Schipper C.A., Lo C.F., Kou G.H.,
Vlak J.M.;

RT "Analysis of a genomic segment of white spot syndrome virus of shrimp
RT containing ribonucleotide reductase genes and repeat regions.";
RL J. Gen. Virol. 81:307-316(2000).
DR EMBL; AF099142; AAF18485.1; -.
KW Hypothetical protein.
SQ SEQUENCE 194 AA; 22267 MW; 73254CF17E9678C CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 61 GDDDDDD 68

RESULT 57
ID 091LE7 PRELIMINARY; PRT; 194 AA.
AC 091LE7;
DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE ORF94.
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
NC NCB1_TaxID=92652;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21342572; PubMed=11448154;
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
Tarchini R., Piers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;

RT "The white spot syndrome virus DNA genome sequence.";
RL Virology 286:7-22(2001).
RN (2)

RP SEQUENCE FROM N.A.
 RA van Hulzen M.C.W., Wittevelde J., Peters S., Kloosterboer N.,
 RA Tarchini R., Fiers W., Sandbink H., Lankhorst R.K., Vlak J.M.,
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF469029; AAK7763.1; -
 SQ SEQUENCE 194 AA; 22281 MW; 19FE26A0504DC78 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 436 GDDDDDD 443
 61 GDDDDDD 68

RESULT 58
 ID OQOT60 PRELIMINARY; PRT; 194 AA.

AC OQOT60; 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE WSSV234.

OS White spot syndrome virus (WSSV).
 OC Viruses, dsDNA viruses, no RNA stage; Nimaviridae.
 OX NCBI_TaxID=92652;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RX MEDLINE=20517548; Pubmed=11062040;

RA Tsai M.F., Yu H.T., Tseng H.F., Lue J.H., Chou C.M., Huang C.J.,
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.,
 RT "Identification and characterization of a shrimp white spot syndrome
 RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
 RT cellular-type thymidine kinase and thymidylate kinase.";
 RL Virology 277:100-110(2000).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RX MEDLINE=21844071; Pubmed=11853398;

RA Chen L.L., Lue J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
 RA Lo C.F., Kou G.H.,
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
 RT spot syndrome virus and characterization of the motif important for
 RT targeting VP35 to the nuclei of transfected insect cells.";
 RL Virology 293:44-53(2002).

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RA Lo C.-F., Kou G.-H.,
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF440570; AAL89102.1; -

SQ SEQUENCE 194 AA; 22253 MW; 733FFC0B214CCD3D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 436 GDDDDDD 443
 61 GDDDDDD 68

Db 61 GDDDDDD 68

RESULT 59
 ID OQIP55 PRELIMINARY; PRT; 198 AA.

AC OQIP55; 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE CG31771-PA.
 GN CG31771.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; Pubmed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhargava D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertys J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegami C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,
 RA Palazzo D.M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).

RN [2]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dreenek D., Fartan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Pargag V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

```

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Flybase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003642; AAN1086.1; -.
SQ SEQUENCE 198 AA; 21562 MW; 30E7CC988146EC17 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 198;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 82 DDDDDDDG 89

RESULT 60
ID Q9C7Y9 PRELIMINARY; PRT; 198 AA.
AC Q9C7Y9;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Hypothetical protein (unknown protein).
GN T2J15.12 OR ATG47970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=1130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehner E., Chan A., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dwar K.,
RA Dunn P., Egu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.U., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millscher J., Miranda M., Nguyen M., Nietman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salsberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteback T., Van Aken S., Vaysberg M., Vysotskaja V.S., Walker M.,
RA Wu D., Yu G., Frazer C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).

RP SEQUENCE FROM N.A.
RA lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bahn U., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,

```

```

RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC051631; AAG51532.1; -.
DR EMBL; AF386997; AAK62442.1; -.
DR EMBL; BP000731; AAN31873.1; -.
KW Hypothetical protein.
SQ SEQUENCE 198 AA; 21847 MW; C0AC434A40FE7818 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDD 441
DB 47 SEGDDDD 54

RESULT 61
ID Q39756 PRELIMINARY; PRT; 204 AA.
AC Q39756;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-UTL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE PKF1 protein (fragment).
GN PKF1.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicolas C., Nicolas G., Rodriguez D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC "- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X57547; CAA6149.1; -.
DR HSSP; P08631; 1AD5.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pk_kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
SQ SEQUENCE 204 AA; 23305 MW; B4C70D228B1077B0 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 204;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VTHRDLS 136
DB 22 VTHRDLS 29

RESULT 62
ID Q9SWT7 PRELIMINARY; PRT; 206 AA.
AC Q9SWT7;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 23, Last annotation update)
DE HVI1PST.
GN HVI1PST.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9324224; PubMed=10393983;
RA Feuille C., Keller B.;
RT "High gene density is conserved at syntenic loci of small and large
RT grass genomes.";
RL EMBL; AF108009; AAD46470.1; -.
SQ SEQUENCE 206 AA; 21440 MW; 3F0BFF6622035942 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 184 GDDDDDD 191

RESULT 63
Q945P0 PRELIMINARY; PRT; 209 AA.
AC Q945P0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transcriptional activator FHA1.
GN FHA1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xanthi;
RA Kim M.I., An U.W., Pai H.S.;
RT "NtFHA1, transcription activator containing forkhead-associated
RT domain.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF41856; AAL05884.1; -.
DR InterPro; IPR000253; FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PSS0006; FHA DOMAIN; 1.
SQ SEQUENCE 209 AA; 22601 MW; E3440C7327070176 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 209;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 444
DB 164 DDDDDDD 171

RESULT 64
Q28960 PRELIMINARY; PRT; 211 AA.
AC Q28960;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein AF1309.
GN AF1309.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Claydon R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kurlavsky A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Kiehlhuth J., Lee N.H., Sutton G.G., Gill S.,
RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne U.D., Weidman J.F., McDonald L., Utterback T.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370 (1997).
DR EMBL; AE001013; AAB89938.1; -.
DR TIGR; AF1309; -.
DR InterPro; IPR000374; Pcrtransf.
DR Pfam; PF01148; CTP_transf_1; 1.
DE Hypothetical protein; Complete proteome.
SQ SEQUENCE 211 AA; 22882 MW; EB258FA5D12855CA CRC64;

Query Match 1.8%; Score 8; DB 17; Length 211;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PLPLPLA 336
DB 174 PLPLPLA 181

RESULT 65
Q91VH4 PRELIMINARY; PRT; 213 AA.
AC Q91VH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 23.6 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014284; AAH14284.1; -.
DR InterPro; IPR002038; Osteopontin.
DR Pfam; PF00865; Osteopontin; 1.
DR PRINTS; PR00216; OSTEOPONTIN.
DE Hypothetical protein.
GN NON TER
FT NON TER
SQ SEQUENCE 213 AA; 23625 MW; 50CC5EDB9E91797 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 213;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 444
DB 7 DDDDDDD 14

RESULT 66
Q8RZR9 PRELIMINARY; PRT; 218 AA.
AC Q8RZR9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

```

DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
 DE B140D12.10 protein.
 GN B140D12.10
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
 RT clone:B140D12.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003537; BAB86162.1; -.
 DR Gramene; Q8RZR9; -.
 SQ SEQUENCE 218 AA; 23468 MW; 9B158986E7938F29 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 218;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443
 DB 109 GDDDDDD 116

RESULT 67

ID 08X4P6 PRELIMINARY; PRT; 229 AA.

AC 08X4P6;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
 DE Putative antiterminator Q of prophage CP-933X (Antitermination
 DE protein).
 GN 21874 OR ECS1620.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=83334;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Groetbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hatori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL; AE005330; AAG55971.1; -.
 DR EMBL; AP002555; BAB35043.1; -.
 KM Complete proteome.
 SQ SEQUENCE 229 AA; 26402 MW; E6ABF81BE9476667 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 229;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 290 TLERLKL 297
 DB 142 TLERLKL 149

RESULT 68

ID 08C215 PRELIMINARY; PRT; 230 AA.

AC 08C215;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Enhancer of zeste homolog 2 (Fragment).
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD; TISSUE=Thymus;
 RX MEDLINE=2234683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK088575; BAC40432.1; -.
 FT NON TER 230
 SQ SEQUENCE 230 AA; 26749 MW; 43706ABA667E9691 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 230;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 444
 DB 174 DDDDDDD 181

RESULT 69

ID 050439 PRELIMINARY; PRT; 237 AA.

AC 050439;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
 DE Tnpr protein.
 GN Tnpr.
 OS Mycobacterium smegmatis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1772;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92078082; PubMed=1660454;
 RA Ciriillo J.D., Barietta R.G., Bloom B.R., Jacobs W.R. Jr.;
 RT "A novel transposon trap for mycobacteria: isolation and
 RT characterization of IS1096.";
 RL J. Bacteriol. 173:7772-7780(1991).
 DR EMBL; M76495; AAA8488.1; -.
 SQ SEQUENCE 237 AA; 26293 MW; B4A0C49525C68EE CRC64;

Query Match 1.8%; Score 8; DB 2; Length 237;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 FDLSEGD 438
 DB 119 FDLSEGD 126

RESULT 70

Q9AS90
ID Q9AS90 PRELIMINARY; PRT; 246 AA.
AC Q9AS90;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE P0028E10.1 protein (P0701D05.18 protein).
GN P0028E10.1 OR P0701D05.18.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
clone:P0028E10.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
clone:P0701D05.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002912; BAB39898.1; -.
DR EMBL; AP003301; BAB64804.1; -.
DR Gramene; Q9AS90; -.
SQ SEQUENCE 246 AA; 25160 MW; F9B39125F8A5A534 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 246;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDD 443
DB 179 GDDDDDD 186

RESULT 71
Q9M1H5
ID Q9M1H5 PRELIMINARY; PRT; 250 AA.
AC Q9M1H5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Hypothetical 29.2 kDa protein.
GN T14E10.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
RA Rüd S., Lemcke K., Meyer K.F.X., Quetler F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138656; CAB7572.1; -.
KM Hypothetical protein.
SQ SEQUENCE 250 AA; 29183 MW; E153440480BBSF6 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 250;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 293 RLKLEKD 300

DB 54 RLKLEKD 61
RESULT 72
Q9GR09
ID Q9GR09 PRELIMINARY; PRT; 258 AA.
AC Q9GR09;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Hypothetical 28.1 kDa protein.
GN L6520.05.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masny D., Fumelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.W., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145 (1998).
DR EMBL; AL46005; CAC14532.1; -.
KM Hypothetical protein.
SQ SEQUENCE 258 AA; 28116 MW; 52A1CB8119F2B079 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 258;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDD 443
DB 70 GDDDDDD 77

RESULT 73
Q9CXN1
ID Q9CXN1 PRELIMINARY; PRT; 259 AA.
AC Q9CXN1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 3110052N05Rik protein.
GN 3110052N05Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koichi H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuke S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK014208; BAB29206.1; -;
 DR MGD; MGI:1924237; 3110052N05R1k.
 DR InterPro; IPR006357; HAD_SF_11A.
 DR InterPro; IPR006355; HAD_SF_11A_hyp2.
 DR InterPro; IPR005834; Hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR TIGRPFAMS; TIGR01460; HAD-SF-11A; 1.
 DR TIGRPFAMS; TIGR01458; HAD-SF-11A-hyp3; 1.
 SQ SEQUENCE 259 AA; 28730 MW; C8E34DF0DC9A8 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 259;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 LERLKLLE 298
 DB 57 LERLKLLE 64

RESULT 74

Q9S1S0 PRELIMINARY; PRT; 260 AA.
 ID Q9S1S0;
 AC Q9S1S0;
 DT 01-MAY-2000 (TrEMBLrel. 13; Created)
 DT 01-MAY-2000 (TrEMBLrel. 13; Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23; Last annotation update)
 DE At2g25220 protein.
 GN At2g25220.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RT Nature 402:761-768(1999).
 RL Nature 402:761-768(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AC007070; AAD23669.1; -;
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR Pfam; PF00669; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 260 AA; 28793 MW; 5A469E1CB1596BEC CRC64;

Query Match 1.8%; Score 8; DB 10; Length 260;
 Best Local Similarity 100.0%; Pred. No. 20;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 129 VIHRDLKS 136
 DB 125 VIHRDLKS 132

RESULT 75

Q819F4 PRELIMINARY; PRT; 269 AA.
 ID Q819F4;
 AC Q819F4;
 DT 01-MAR-2003 (TrEMBLrel. 23; Created)
 DT 01-MAR-2003 (TrEMBLrel. 23; Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23; Last annotation update)
 DE Aspartic acid-rich protein.
 GN ARP.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Dobson S., Kumar R., Bracchi-Ricard V., Al-Muriani S.W.K., Damuni Z.,
 RA Chakrabarti D., Barik S.;
 RT "Characterization of a unique aspartate-rich protein (ARP) of the
 RT SET/PAF-family in the human malaria parasite, Plasmodium falciparum,
 RT which inhibits protein phosphatase 2A";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF543685; AA016228.1; -;
 SQ SEQUENCE 269 AA; 31710 MW; 84EA2D28378E243F CRC64;

Query Match 1.8%; Score 8; DB 5; Length 269;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 DB 242 DDDDDDDG 249

Search completed: December 5, 2003, 09:31:40
 Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:21:02 ; Search time 46 Seconds
(without alignments)
1570.013 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSLSASFQIKFDLQFFE.....GDDDDDDDDGGERNDNMNDSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_19Jun03:*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	455	100.0	455	21	AA818657
2	455	100.0	455	21	AA83278
3	455	100.0	455	21	AA84331
4	447	98.2	473	22	AA85322
5	331	72.7	800	22	AA871957
6	331	72.7	800	23	AA856573
7	331	72.7	800	23	ABP43736
8	269	59.1	349	22	AA875571
9	124	27.3	124	22	AA85552

10	118	25.9	141	21	AA803583	Human secreted pro
11	80	17.6	92	22	AA87295	Novel central nerv
12	80	17.6	92	22	AA87295	Novel signal trans
13	56	12.3	144	22	AB806092	Novel human diagno
14	15	3.1	14	21	AA83280	Human survival reg
15	15	3.1	13	22	AA871969	Human TGF-beta rec
16	11	2.4	977	12	AB871694	Drosophila melanog
17	10	2.2	598	13	AA820470	Soluble ysc alpha.
18	10	2.2	729	11	AA805457	KEK1 amino acid se
19	10	2.2	764	21	AA801897	Nonsense-mediated
20	10	2.2	764	21	AA801897	Nonsense-mediated
21	10	2.2	1089	17	AA801896	Yeast NMD2 carboxy
22	10	2.2	1089	20	AA801896	Nonsense-mediated
23	10	2.2	1089	20	AA801896	Yeast Nmd2p involv
24	9	2.0	133	21	AA821315	Yeast Nmd2p involv
25	9	2.0	156	21	AA821314	Arabidopsis thalia
26	9	2.0	201	21	AA807429	Arabidopsis thalia
27	9	2.0	224	21	AA807428	Arabidopsis thalia
28	9	2.0	224	24	AA808115	Human kinase and p
29	9	2.0	502	15	AA85371	Mouse Activin rece
30	9	2.0	502	20	AA83304	Human MALK-1 clone
31	9	2.0	505	14	AA841922	MISR3. Ratius rat
32	9	2.0	505	17	AA803761	Mullerian inhibiti
33	9	2.0	648	23	ABP30294	Streptococcus poly
34	9	2.0	652	23	ABP27599	Streptococcus poly
35	9	2.0	652	23	ABP27600	Streptococcus poly
36	9	2.0	656	23	ABP16259	Human kinase PKIN-
37	9	2.0	909	22	AA803554	Human protein kina
38	9	2.0	911	24	ABU11053	Human protein NOVI
39	9	2.0	915	24	ABU11053	Human protein NOVI
40	9	2.0	2391	15	AA85694	Carbamoyl-phosphat
41	9	2.0	2527	24	AA808105	Human kinase and p
42	8	1.8	20	22	AA883387	Lunasin fragment #
43	8	1.8	20	22	AA862621	Soybean lunasin pe
44	8	1.8	21	22	AA883386	Lunasin fragment #
45	8	1.8	21	22	AA883393	Lunasin related pe
46	8	1.8	21	22	AA862620	Soybean lunasin pe
47	8	1.8	22	22	AB850623	Human liver peptid
48	8	1.8	22	22	AA883385	Lunasin fragment #
49	8	1.8	22	22	AB830607	Peptide #3258 enco
50	8	1.8	22	22	AB835771	Peptide #3277 enco
51	8	1.8	22	22	AB821195	Peptide #3194 enco
52	8	1.8	22	22	AA856580	Human brain expres
53	8	1.8	22	22	AA868960	Human bone marrow
54	8	1.8	22	22	AA816787	Peptide #2221 enco
55	8	1.8	22	22	AA830607	Peptide #3308 enco
56	8	1.8	22	22	AA804504	Peptide #3186 enco
57	8	1.8	22	22	AA862619	Soybean lunasin pe
58	8	1.8	22	23	AB838546	Human peptide enco
59	8	1.8	25	22	AA883395	Lunasin related pe
60	8	1.8	31	22	AB849409	Human liver peptid
61	8	1.8	31	22	AB829408	Peptide #2059 enco
62	8	1.8	31	22	AB834590	Peptide #2096 enco
63	8	1.8	31	22	AB819996	Protein #1995 enco
64	8	1.8	31	22	AA855379	Human brain expres
65	8	1.8	31	22	AA867774	Human bone marrow
66	8	1.8	31	22	AA815583	Peptide #2017 enco
67	8	1.8	31	22	AA828075	Peptide #2112 enco
68	8	1.8	31	22	AA803327	Peptide #2009 enco
69	8	1.8	31	23	AB837322	Human peptide enco
70	8	1.8	41	22	AA883381	Lunasin fragment #
71	8	1.8	41	22	AA862615	Soybean lunasin pe
72	8	1.8	42	22	AB851319	Human liver peptid
73	8	1.8	42	22	AA883380	Lunasin fragment #
74	8	1.8	42	22	AB831264	Peptide #3915 enco
75	8	1.8	42	22	AB836472	Peptide #3978 enco
76	8	1.8	42	22	AB821815	Protein #3814 enco
77	8	1.8	42	22	AA857237	Human brain expres
78	8	1.8	42	22	AA869636	Human bone marrow
79	8	1.8	42	22	AA817452	Peptide #3886 enco
80	8	1.8	42	22	AA829972	Peptide #4009 enco
81	8	1.8	42	22	AA805129	Peptide #3811 enco
82	8	1.8	42	22	AA862614	Soybean lunasin pe


```

QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSNRSEMDMDHIMTATDVAKGMY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSNRSEMDMDHIMTATDVAKGMY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180
QY 181 LPVSECTDYTSYGIVLWEMLTREVPRKGLGLOVAMLVKERNRLTTPSSCPSPFALLH 240
DB 181 LPVSECTDYTSYGIVLWEMLTREVPRKGLGLOVAMLVKERNRLTTPSSCPSPFALLH 240
QY 241 QCWEADAKKRPSPFQKIIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLEKLEKLE 300
DB 241 QCWEADAKKRPSPFQKIIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLEKLEKLE 300
QY 301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360
DB 301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360
QY 361 TATSNEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMOINMOAKONSSKTTSTRRGK 420
DB 361 TATSNEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMOINMOAKONSSKTTSTRRGK 420
QY 421 KVMNALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455
DB 421 KVMNALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455

```

RESULT 2

AA83278 standard; Protein; 455 AA.

AA83278;

16-AUG-2000 (first entry)

Human survival regulating kinase (SRK).

Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;
HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
apoptosis; cell survival; nuclear targeting; tumour; human;
autoimmune disease.

Homo sapiens.

WO200022142-A2.

20-APR-2000.

20-SEP-1999; 99WO-US22008.

13-OCT-1998; 98US-0104088.

(ONYX-) ONYX PHARM INC.

Rugieri R, Callow M, Diaz P,

WPI; 2000-317994/27.

N-PDB; AA293783.

Novel human survival regulating kinase polypeptide for screening agents

which modulate biological pathways associated with SRK useful in

treating autoimmune diseases, tumors and apoptosis-related disorders

Claim 4; Figure 2; 62pp; English.

Survival regulating kinases (SRK) are a class of proteins involved in
cell signal transduction pathways such as mitogen-activated protein
kinase pathways. A protein kinase activity means that the SRK can
catalyse a reaction in which a phosphate group is transferred from a
phosphate donor to a phosphate acceptor amino acid residue,

CC preferably the hydroxyl side chain of a serine or threonine.
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
CC activity is similar to that of a MAPKK such as Raf. has a range of
CC other activities including a cell growth-regulatory activity, a cell
CC survival promoting activity, a HAX-1 binding activity, an apoptosis
CC suppressing activity, a MAPK activation or stimulatory activity, a
CC nuclear targeting activity and a SRK-specific immunogenic activity.
CC SRK is useful for identifying agents which modulate cellular
CC transformations mediated by Ras and SRK and agents that modulate the
CC apoptosis suppression activity of SRK. This information may be useful
CC in the treatment of autoimmune diseases, tumours and apoptosis
CC related disorders.

Sequence 455 AA;

Query Match 100.0%; Score 455; DB 21; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSLSGASFVQIRFDLQFPENCGGGSGSVYAKWISODKEVAVKKLTKKEAEILSVL 60
DB 1 MSLSGASFVQIRFDLQFPENCGGGSGSVYAKWISODKEVAVKKLTKKEAEILSVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSNRSEMDMDHIMTATDVAKGMY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSNRSEMDMDHIMTATDVAKGMY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180
QY 181 LPVSECTDYTSYGIVLWEMLTREVPRKGLGLOVAMLVKERNRLTTPSSCPSPFALLH 240
DB 181 LPVSECTDYTSYGIVLWEMLTREVPRKGLGLOVAMLVKERNRLTTPSSCPSPFALLH 240
QY 241 QCWEADAKKRPSPFQKIIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLEKLEKLE 300
DB 241 QCWEADAKKRPSPFQKIIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLEKLEKLE 300
QY 301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360
DB 301 LSFKEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360
QY 361 TATSNEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMOINMOAKONSSKTTSTRRGK 420
DB 361 TATSNEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMOINMOAKONSSKTTSTRRGK 420
QY 421 KVMNALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455
DB 421 KVMNALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455

```

RESULT 3

AA84321 standard; Protein; 455 AA.

AA84321;

12-JUL-2000 (first entry)

A human cardiovascular system associated protein kinase-2.

Human; cardiovascular system associated protein kinase-2; CSAPK-2;

signalling pathway; cell growth; cell differentiation; gene mapping;

tissue typing; forensic identification; cardiovascular disease;

congestive heart failure; transgenic animal.

Homo sapiens.

WO200014212-A1.

16-MAR-2000.

```

PF 09-SEP-1999; 99MO-US20631.
XX
XX 09-SEP-1998; 98US-0099657.
PR 29-SEP-1998; 98US-0163115.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Acton S;
XX
XX WPI; 2000-271053/23.
DR N-PSDB; AAZ299726, AAZ299727.
XX
XX New nucleic acid encoding cardiovascular system associated protein
PT kinase, used e.g. for diagnosis, treatment and prevention of
PT cardiovascular disease
XX
XX Claim 2; Fig 2; 163pp; English.
XX
XX The present sequence represents a human cardiovascular system associated
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling
CC pathways associated with cell growth and differentiation. The CSAPK
CC polypeptides and polynucleotides are used to screen for agents that
CC specifically modulate CSAPK, which are potential therapeutic agents.
CC They are also used for diagnosis, prognosis or monitoring of
CC CSAPK-related diseases. Gene mapping, tissue typing and forensic
CC identification, and for treating or preventing disorders associated
CC with aberrant CSAPK expression or activity, especially cardiovascular
CC diseases such as congestive heart failure. They can also be used in
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate
CC transgenic animals.
XX
XX Sequence 455 AA;
SQ
Query Match 100.0%; Score 455; DB 21; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSLGASFNVOIKRFDLQFFENCGGSGSVYRAKWTISQDEVAVKKLKTEKAEITLSVL 60
DB 1 MSSLGASFNVOIKRFDLQFFENCGGSGSVYRAKWTISQDEVAVKKLKTEKAEITLSVL 60
QY 61 SHNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
DB 61 SHNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDIKSRNVIIADGVLIKIDFGASRPHNHTTMSLVGTFPMNAPEVIQS 180
DB 121 LHMEAPVKVIHRDIKSRNVIIADGVLIKIDFGASRPHNHTTMSLVGTFPMNAPEVIQS 180
QY 181 LPVSECTDYSYGVVLEMLTREVPFKGLEGLQAVAVVKNRRLTPSCCPSPFALLH 240
DB 181 LPVSECTDYSYGVVLEMLTREVPFKGLEGLQAVAVVKNRRLTPSCCPSPFALLH 240
QY 241 QCWEADAKKRPSPFOKISILESMSNDLSLPDKCNFLHNKAEMRCIEATLETLKLERD 300
DB 241 QCWEADAKKRPSPFOKISILESMSNDLSLPDKCNFLHNKAEMRCIEATLETLKLERD 300
QY 301 QCEHEADAKKRPSPFOKISILESMSNDLSLPDKCNFLHNKAEMRCIEATLETLKLERD 360
DB 301 QCEHEADAKKRPSPFOKISILESMSNDLSLPDKCNFLHNKAEMRCIEATLETLKLERD 360
QY 361 TATSNGGCHMNPISLOAMLMGFGDIFSNMKAAGVNHSGMIOIMOKAKNSKTTXRRGK 420
DB 361 TATSNGGCHMNPISLOAMLMGFGDIFSNMKAAGVNHSGMIOIMOKAKNSKTTXRRGK 420
QY 421 KVNNALGSPDFDISEGDDDDDCGEEDNDMDNSE 455
DB 421 KVNNALGSPDFDISEGDDDDDCGEEDNDMDNSE 455

```

```

XX
XX AAM25322;
AC
XX 16-OCT-2001 (First entry)
DT
XX
XX Human protein sequence SEQ ID NO:837.
DE
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; anticancer; osteopathic; eczema;
KW dermatological; antiallergic; antiashtmatic; antidiabetic; cyostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
XX Homo sapiens.
OS
XX
XX MO200153455-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 22-DEC-2000; 2000MO-US35017.
PF
XX
XX 23-DEC-1999; 99US-0471275.
PR 21-JUN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dmanac RT;
PI
XX
XX WPI; 2001-457603/49.
DR N-PSDB; AAH99263.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
PT
XX
XX Claim 20; Page 191; 1217pp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antineoplastic; antirheumatic;
CC antidiabetic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC anticancer; osteopathic; dermatological; antiallergic; antiashtmatic;
CC antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmune; genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
XX Sequence 473 AA;
SQ
Query Match 98.2%; Score 447; DB 22; Length 473;
Best Local Similarity 100.0%; Pred. No. 0;

```

Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGASFOVQIKFDLQFFENCGGSGFSGVYRAKWIISQKEVAVVKLLKIEKEAETLSVL 60
 DB 19 MSLSGASFOVQIKFDLQFFENCGGSGFSGVYRAKWIISQKEVAVVKLLKIEKEAETLSVL 78
 QY 61 SHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
 DB 79 SHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 138
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIGS 180
 DB 139 LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIGS 198
 QY 181 LVSSECTDITYSGVVLWEMLTREVPKGLGQVAMLVYKNERLITPSSCRSPFALLH 240
 DB 199 LVSSECTDITYSGVVLWEMLTREVPKGLGQVAMLVYKNERLITPSSCRSPFALLH 258
 QY 241 QCMEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKLKLERD 300
 DB 259 QCMEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKLKLERD 318
 QY 301 LSFKEQELKERERRLKMEQKLTQSNTPPLLPLAARMSEESYFESKTEESNSAEMSCQI 360
 DB 319 LSFKEQELKERERRLKMEQKLTQSNTPPLLPLAARMSEESYFESKTEESNSAEMSCQI 378
 QY 361 TATSNGEGHGMPSLOAMMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRRGX 420
 DB 379 TATSNGEGHGMPSLOAMMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRRGX 438
 QY 421 KYNMALGFSDPDLSEGDGDDDDGDEE 447
 DB 439 KYNMALGFSDPDLSEGDGDDDDGDEE 465

RESULT 5
 AAB71957
 ID AAB71957 standard; protein; 800 AA.
 AC AAB71957;
 XX
 DT 11-MAY-2001 (first entry)
 XX
 DE Human TGF-beta receptor encoded by cDNA clone HDPSM48.
 XX
 KW Human; anti-sclerotic; dermatological; immunosuppressive; cytostatic;
 KW anti-inflammatory; anti-HIV; immunostimulant; cardiant; vascular;
 KW ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;
 KW antiparkinsonian; antimicrobial; vulnerary; gene therapy; infection;
 KW transforming growth factor; TGF-beta receptor; immune disorder;
 KW hyperproliferative disorder; cardiovascular disease; angiogenesis;
 KW neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN MO200112670-A1.
 XX
 PD 22-FEB-2001.
 XX
 PF 10-AUG-2000; 2000MO-US21736.
 XX
 PR 13-AUG-1999; 99US-0148682.
 XX
 PR 20-SEP-1999; 99US-0154887.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J;
 XX
 DR WPI; 2001-202858/20.
 XX
 DR N-PSDB; AAF75336.
 XX
 PT Nucleic acid molecules encoding 12 transforming growth factor-beta
 PT receptor polypeptides, useful for preventing, diagnosing and treating

PT e.g. cancers, Parkinson's disease and diabetic retinopathy -
 XX
 XX Claim 11; Page 293-295; 311pp; English.
 XX
 CC The present sequence is one of 12 novel human transforming growth factor
 CC (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides
 CC and polypeptides may be used in the prevention, diagnosis and treatment
 CC of diseases associated with inappropriate polypeptide expression. Such
 CC diseases include immune disorders (e.g. multiple sclerosis, systemic
 CC lupus erythematosus and human immunodeficiency virus (HIV) infections),
 CC hyperproliferative disorders (e.g. cancers and Gaucher's disease),
 CC cardiovascular diseases (e.g. Scimitar syndrome, Chagas' cardiomyopathy
 CC and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)
 CC and infectious diseases. The polynucleotides and polypeptides are also
 CC useful for promoting wound healing, regeneration and/or chemotaxis. The
 CC polynucleotides and their complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples. The polypeptides may be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of protein expression and activity. The anti-TGF-beta receptor
 CC antibodies may be used to down regulate expression and activity and as
 CC diagnostic agents for detecting the presence of the polypeptides in
 CC samples.
 CC
 SQ Sequence 800 AA;
 Query Match 72.7%; Score 331; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 3.3e-304;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGASFOVQIKFDLQFFENCGGSGFSGVYRAKWIISQKEVAVVKLLKIEKEAETLSVL 60
 DB 1 MSLSGASFOVQIKFDLQFFENCGGSGFSGVYRAKWIISQKEVAVVKLLKIEKEAETLSVL 60
 QY 61 SHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
 DB 61 SHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIGS 180
 DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIGS 180
 QY 181 LVSSECTDITYSGVVLWEMLTREVPKGLGQVAMLVYKNERLITPSSCRSPFALLH 240
 DB 181 LVSSECTDITYSGVVLWEMLTREVPKGLGQVAMLVYKNERLITPSSCRSPFALLH 240
 QY 241 QCMEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKLKLERD 300
 DB 241 QCMEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKLKLERD 300
 QY 301 LSFKEQELKERERRLKMEQKLTQSNTPPLL 331
 DB 301 LSFKEQELKERERRLKMEQKLTQSNTPPLL 331

RESULT 6
 AAB65673
 ID AAB65673 standard; Protein; 800 AA.
 AC AAB65673;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 201.
 XX
 KW Human; mouse; protein kinase; antiarthritic; anti-sclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; anti-inflammatory; antiasclerotic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.


```

Db      121 LHMEAPVKVIHRDLKSRNVIIAADGVLKICDGFASRFHNHTTHMSLVGTFFPMMAPEVIQS 180
QY      181 LPVSECTDYTSYGVLWEMLTREVPFKGLEQVAMLVVEKNERLTIIPSSCRSPFAELH 240
Db      181 LPVSECTDYTSYGVLWEMLTREVPFKGLEQVAMLVVEKNERLTIIPSSCRSPFAELH 240
QY      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSFLHNKAEMRCIEATLERLKLERD 300
Db      241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSFLHNKAEMRCIEATLERLKLERD 300
QY      301 LSFKEQELKEREERRLKMEOKLTQSGNTPL 331
Db      301 LSFKEQELKEREERRLKMEOKLTQSGNTPL 331

RESULT 8
AAG75571
ID      AAG75571 standard; Protein; 349 AA.
XX
XX      AAG75571;
AC
XX      03-SEP-2001 (first entry)
DT
XX      Human colon cancer antigen protein SEQ ID NO:6335.
DE
XX      Human colon cancer antigen; colon cancer antigen; diagnosis; detection;
KM      colorectal carcinoma.
XX
XX      Homo sapiens.
OS
XX      WO200122920-A2.
PN
XX      05-APR-2001.
PD
XX      28-SEP-2000; 2000WO-US26524.
PF
XX      29-SEP-1999; 99US-0157137.
PR      03-NOV-1999; 99US-0163280.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
PA
XX      Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX      WPI: 2001-235357/24.
XX      N-PSDB; AAH34976.
DR
XX
XX      Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT      useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX      Claim 11; Page 7789-7790; 9803pp; English.
XX
XX      AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC      cancer-associated nucleic acid molecules (N) and proteins (P), where
CC      the proteins are collectively known as colon cancer antigens. The colon
CC      cancer antigens have cytostatic activity and can be used in gene
CC      therapy and vaccine production. N and P may be used in the prevention,
CC      diagnosis and treatment of diseases associated with inappropriate P
CC      expression. For example, N and P may be used to treat disorders
CC      associated with decreased expression by rectifying mutations or deletions
CC      in a patient's genome that affect the activity of P by expressing of P.
CC      Inactive proteins or to supplement the patients own production of P.
CC      Additionally, N may be used to produce the colon cancer-associated P,
CC      by inserting the nucleic acids into a host cell and culturing the cell
CC      to express the proteins. N and P can be used in the prevention, diagnosis
CC      and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC      and AAG77789 represent sequences used in the exemplification of the
CC      present invention.
CC      N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC      missing at time of publication, meaning no sequences are present for
CC      SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX      Sequence 349 AA;

```

```

Query Match      59.1%; Score 269; DB 22; Length 349;
Best Local Similarity 100.0%; Pred. No. 7,4e-246;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSSLGASPVQIKFDDLQFFENCSSFGSVYRAKMIQDKEVAVKKLKIEKEAETLSVL 60
Db      57 MSSLGASPVQIKFDDLQFFENCSSFGSVYRAKMIQDKEVAVKKLKIEKEAETLSVL 116
QY      61 SHRNIIOFPYGVLEPPNNGIYEVASLSGLYIYINSNSSEEDMDHIMWADVAKGMHY 120
Db      117 SHRNIIOFPYGVLEPPNNGIYEVASLSGLYIYINSNSSEEDMDHIMWADVAKGMHY 176
QY      121 LHMEAPVKVIHRDLKSRNVIIAADGVLKICDGFASRFHNHTTHMSLVGTFFPMMAPEVIQS 180
Db      177 LHMEAPVKVIHRDLKSRNVIIAADGVLKICDGFASRFHNHTTHMSLVGTFFPMMAPEVIQS 236
QY      181 LPVSECTDYTSYGVLWEMLTREVPFKGLEQVAMLVVEKNERLTIIPSSCRSPFAELH 240
Db      237 LPVSECTDYTSYGVLWEMLTREVPFKGLEQVAMLVVEKNERLTIIPSSCRSPFAELH 296
QY      241 QCWEADAKKRPSPFKQIISILESMSNDTSL 269
Db      297 QCWEADAKKRPSPFKQIISILESMSNDTSL 325

RESULT 9
AAB65552
ID      AAB65552 standard; Protein; 124 AA.
XX
XX      AAB65552;
AC
XX      27-MAR-2001 (first entry)
DT
XX      C-terminus specific to novel human protein kinase MLK4B.
DE
XX      Human; protein kinase; antiarthritic; antisclerotic; immunosuppressive;
KM      cardiac; renal; antiinflammatory; antiasthmatic; osteoparathic;
KM      dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KM      immune disorder; cardiovascular disease; neurodegenerative disease;
KM      cancer; autoimmune disorder; stroke; inflammatory bowel disease;
XX      inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
XX      Homo sapiens.
OS
XX      WO200073469-A2.
PN
XX      07-DEC-2000.
PD
XX      26-MAY-2000; 2000WO-US14842.
PF
XX      28-MAY-1999; 99US-0136503.
PR
XX      (SUGR-) SUGEN INC.
PA
XX      Plowman GD, Martinez R, Whyte D, Sudersanam S;
PI
XX      WPI: 2001-032161/04.
DR
XX
XX      Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT      treating immune-related diseases and disorders, cardiovascular disease,
XX      neurodegenerative diseases and/or cancers -
XX
XX      Example 1; Page 116; 310pp; English.
XX
XX      The present sequence is given in a specification relating to novel
CC      protein kinases. The protein kinases and the nucleic acids that encode
CC      them may be used in the treatment and diagnosis of diseases
CC      associated with inappropriate kinase expression such as immune-related
CC      diseases and disorders, cardiovascular disease, neurodegenerative
CC      diseases and/or cancers. The nucleic acids and complementary sequences
CC      may also be used as DNA probes in diagnostic assays. The kinase
CC      polypeptides may be used as antigens in the production of antibodies of
CC      kinase expression and activity. Anti-kinase antibodies and kinase

```

CC antagonists may also be used to down regulate kinase expression and
CC activity. Diseases related to kinase expression and activity include
CC rheumatoid arthritis, atherosclerosis, autoimmune disorders,
CC complications of organ transplantation, myocardial infarction, immune
CC disorders, cardiomyopathies, strokes, renal failure, oxidative-stress
CC related disorders, chronic inflammatory bowel disease, chronic
CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
CC disorders.

XX Sequence 124 AA;

Query Match 27.3%; Score 124; DB 22; Length 124;

Best Local Similarity 100.0%; Pred. No. 4.7e-109; Mismatches 0; Gaps 0;
Matches 124; Conservative 0; Indels 0; Gaps 0;

QY 332 LPLAARNSSESYFESKTEESNSAEMSCQITATSNBEGHGNPISQAMLMGFQDIPSMNK 391
DB 1 LPLAARNSSESYFESKTEESNSAEMSCQITATSNBEGHGNPISQAMLMGFQDIPSMNK 60

QY 392 AGAVMHSGMGINQAKNNSKRTSKRGKVMNALGFSDDLSRGDDDDDDGDEEENDM 451
DB 61 AGAVMHSGMGINQAKNNSKRTSKRGKVMNALGFSDDLSRGDDDDDDGDEEENDM 120

QY 452 DNSE 455
DB 121 DNSE 124

RESULT 10
AAG03583

ID AAG03583 standard; Protein: 141 AA.

XX AAG03583;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 7664.

XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; Chromosome mapping.

XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI, 2000-500381/45.

XX N-PSDB; AAC03589.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 13; SEQ ID 7664; 71bp + CD-ROM, English.

XX The present sequence is a polypeptide encoded by one of a large number
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX were prepared from total human RNAs or polyA+ RNAs derived from 30
XX different tissues. EST sequences usually correspond mainly to the 3'
XX untranslated region (UTR) of the mRNA because they are often obtained
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX those cases where longer cDNA sequences have been obtained, the full 5'

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 141 AA;

Query Match 25.9%; Score 118; DB 21; Length 141;

Best Local Similarity 100.0%; Pred. No. 2.5e-103; Mismatches 0; Gaps 0;
Matches 118; Conservative 0; Indels 0; Gaps 0;

QY 102 MMDHMTATDVAKGMHYLHMEAPYKVIHRDLKSRNVVIAADGVLCDFGASRPHNT 161
DB 1 MMDHMTATDVAKGMHYLHMEAPYKVIHRDLKSRNVVIAADGVLCDFGASRPHNT 60

QY 162 THMSLVGTFPPMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFEGLEGLOVAMLV 219
DB 61 THMSLVGTFPPMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFEGLEGLOVAMLV 118

RESULT 11

ID AAU87295 standard; Protein: 92 AA.

XX AAU87295;

XX 05-JUN-2002 (first entry)

XX Novel central nervous system protein #205.

XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
XX hyperproliferative disorder; neoplasm; cardiovascular disorder;
XX cardiac arrest; cerebrovascular disorder; ischemia; angioneurosis;
XX nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
XX acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
XX adenocarcinoma; reproductive system disorder; testicular feminisation;
XX endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
XX respiratory disorder; renal disorder; kidney failure; blood disorder;
XX myocardial infarction; wound healing; cell proliferation; skin aging;
XX food additive; food preservative; gene therapy.

XX Homo sapiens.

XX WO200155318-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01332.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234977.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236127.
PR 29-SEP-2000; 2000US-0236167.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0248207.
PR 17-NOV-2000; 2000US-0248208.
PR 17-NOV-2000; 2000US-0248209.
PR 17-NOV-2000; 2000US-0248210.
PR 17-NOV-2000; 2000US-0248211.
PR 17-NOV-2000; 2000US-0248212.
PR 17-NOV-2000; 2000US-0248213.
PR 17-NOV-2000; 2000US-0248214.
PR 17-NOV-2000; 2000US-0248215.
PR 17-NOV-2000; 2000US-0248216.
PR 17-NOV-2000; 2000US-0248217.
PR 17-NOV-2000; 2000US-0248218.
PR 17-NOV-2000; 2000US-0248244.
PR 17-NOV-2000; 2000US-0248245.
PR 17-NOV-2000; 2000US-0248264.
PR 17-NOV-2000; 2000US-0248265.
PR 17-NOV-2000; 2000US-0248297.
PR 17-NOV-2000; 2000US-0248299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0255719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
XX
PA Rosen CA, Barash SC, Ruben SM;
XX
PI WPI; 2001-581633/65.
XX
DR N-PSDB; ABR43625.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
XX preventing, treating or ameliorating medical conditions and used as
XX food additives or preservatives -
XX
XX
PS Claim 9; SEQ ID No 813; 837pp; English.
XX

CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. Rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to

CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 17.6%; Score 80; DB 22; Length 92;

Best Local Similarity 100.0%; Pred. No. 1,5e-67; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

QY 54 AELLVLSHNNITQFVGVILLEPPNYGIVTEYASLSGYDYNINRSEMDMDHMTWATD 113
Db 13 AELLVLSHNNITQFVGVILLEPPNYGIVTEYASLSGYDYNINRSEMDMDHMTWATD 72

QY 114 VAKGMHYLHMEAPYKVTHRD 133
Db 73 VAKGMHYLHMEAPYKVTHRD 92

RESULT 12

AAU17234 standard; Protein; 92 AA.

AAU17234;

07-NOV-2001 (first entry)

Novel signal transduction pathway protein, Seq ID 799.

Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;
antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
immune system disorder; rheumatoid arthritis; inflammatory condition;
organ transplant rejection; infection; hepatitis C; blood disorder;
sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
reproductive system; gastrointestinal; liver disorder; AIDS;
acquired immune deficiency syndrome.

Homo sapiens.

MO200154733-A1.

02-AUG-2001.

17-JAN-2001; 2001WO-US01312.

31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251038.
 PR 05-DEC-2000; 2000US-0251867.
 PR 06-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465460/50.
 N-PSDB; AAS27151.

Novel polypeptides useful for diagnosing, treating, preventing and/or
 prognosing disorders related to the proteins, including cancers, immune
 disorders and neuronal disorders -

Claim 1; SEQ ID No 799; 880pp; English.

The invention relates to novel isolated polypeptides (I), and
 polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 diagnosing, preventing and treating diseases including immune system
 disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 transplant rejections, and graft versus host disease, infectious diseases
 (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 other blood-related disorders (sickle cell anaemia), myeloproliferative
 disorders, primary haematopoietic disorders, hyperproliferative
 disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 abnormalities (Down syndrome), ischemic injury (e.g. stroke), renal
 disorders (e.g. glomerulonephritis), cardiovascular disorders
 (e.g. arhythmia), respiratory disorders, dermatological disorders, in
 wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 Addison's disease), reproductive system disorders, gastrointestinal
 disorder (inflammatory disorders), liver disorders (cirrhosis),
 as stimulators of B-cell responsiveness to pathogens, activators of
 T-cells, to induce higher affinity antibodies, and as a means to induce
 tumour proliferation in pathologies e.g. acquired immune deficiency
 syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 pathway protein, amino acid sequences of the invention.

XX

Query Match 17.6%; Score 80; DB 22; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1,5e-67;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

54 AETLSVLSHRNTIOFGVILEPPNYGIVTEYASLSGLYDIYNSNSEMDMDHIMTWARD 113
 |||||
 13 AETLSVLSHRNTIOFGVILEPPNYGIVTEYASLSGLYDIYNSNSEMDMDHIMTWARD 72

Db

QY 114 VAKGMHYLHMEAPVKYIHRD 133
 |||||
 Db 73 VAKGMHYLHMEAPVKYIHRD 92

RESULT 13
 ABG06092

ID ABG06092 standard; Protein; 144 AA.

XX

ABG06092;

AC

13-FEB-2002 (first entry)

DT

Novel human diagnostic protein #6083.

DE

Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

KW

Homo sapiens.

XX

WO200175067-A2.

XX

11-OCT-2001.

XX

30-MAR-2001; 2001WO-US08631.

XX

31-MAR-2000; 2000US-0540217.

PR

23-AUG-2000; 2000US-0649167.

XX

(HYSE-) HYSEQ INC.

PA

Drmancic RT, Liu C, Tang YT;

XX

WPI; 2001-639362/73.

DR

N-PSDB; AAS70279.

XX

New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other, traits and to assess
 biodiversity -

Claim 20; SEQ ID No 36451; 103pp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful in medical
 imaging of sites expressing (II). (I) and (II) are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human
 diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO

CC at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 144 AA;

Query Match 12.3%; Score 56; DB 22; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.2e-44;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVIAAGVAKICDFGSRFHHHTTHMSLNGTFPMMAPEVIOSLPVSEKCDTYSYGV 194
DB 58 VVIAAGVAKICDFGSRFHHHTTHMSLVGTFPMMAPEVIOSLPVSEKCDTYSYGV 113

RESULT 14

ID AAY83280 standard; Protein; 14 AA.

XX AAY83280;

DT 16-AUG-2000 (first entry)

DE Human survival regulating kinase (SRK) antigenic peptide.

KM Survival regulating kinase; SRK; Raf; MAPK; MAPKK; MAPKKK; MAPKKK;
KM HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
KM apoptosis; cell survival; nuclear targeting; tumour; human;
KM autoimmune disease.

OS Homo sapiens.

PN WO200022142-A2.

PD 20-APR-2000.

PF 20-SEP-1999; 99WO-US22008.

PR 13-OCT-1998; 98US-0104088.

XX (ONYX-) ONYX PHARM INC.

PI Ruggieri R, Callow M, Diaz P;

DR WPI; 2000-317994/27.

PT Novel human survival regulating kinase polypeptide for screening agents
PT which modulate biological pathways associated with SRK useful in
PT treating autoimmune diseases, tumors and apoptosis-related disorders
XX
PS Claim 53; Page 6; 62pp; English.

CC Survival regulating kinases (SRK) are a class of proteins involved in
CC cell signal transduction pathways such as mitogen-activated protein
CC kinase pathways. A protein kinase activity means that the SRK can
CC catalyze a reaction in which a phosphate group is transferred from a
CC phosphate donor to a phosphate acceptor amino acid residue.
CC preferably the hydroxyl side chain of a serine or threonine.
CC Substrates for SRK include SRK, MAPK and BAD and SRK's protein kinase
CC activity is similar to that of a MAPKKK such as Raf. has a range of
CC other activities including a cell growth-regulatory activity, a cell
CC survival promoting activity, a HAX-1 binding activity, an apoptosis
CC suppressing activity, a MAPKK activation or stimulatory activity, a
CC nuclear targeting activity and a SRK-specific immunogenic activity.
CC SRK is useful for identifying agents which modulate cellular
CC transformations mediated by Ras and SRK and agents that modulate the
CC apoptosis suppressing activity of SRK. This information may be useful
CC in the treatment of autoimmune diseases, tumours and apoptosis
CC related disorders.
XX
SQ Sequence 14 AA;

Query Match 3.1%; Score 14; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 AKONSKTTSKRRG 419
DB 1 AKONSKTTSKRRG 14

RESULT 15

ID AAB71969 standard; Peptide; 13 AA.

XX AAB71969;

DT 11-MAY-2001 (first entry)

DE Human TGF-beta receptor serine/threonine kinase domain #2.

KM Human; antisclerotic; dermatological; immunosuppressive; cytostatic;
KM antiinflammatory; anti-HIV; immunostimulant; cardiac; vascular;
KM ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;
KM antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;
KM transforming growth factor; TGF-beta receptor; immune disorder;
KM hyperproliferative disorder; cardiovascular disease; angiogenesis;
KM neurological disorder.

OS Homo sapiens.

PN WO200112670-A1.

PD 22-FEB-2001.

PF 10-AUG-2000; 2000WO-US21736.

PR 13-AUG-1999; 99US-0148682.

XX 20-SEP-1999; 99US-0154887.

XX (HDMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J;

DR WPI; 2001-202858/20.

PT Nucleic acid molecules encoding 12 transforming growth factor-beta
PT receptor polypeptides, useful for preventing, diagnosing and treating
PT e.g. cancers, Parkinson's disease and diabetic retinopathy -
XX
PS Disclosure; Page 308; 311pp; English.

CC The present sequence is provided in a specification relating to nucleic
CC acid molecules encoding 12 novel human transforming growth factor
CC (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides
CC and polypeptides may be used in the prevention, diagnosis and treatment
CC of diseases associated with inappropriate polypeptide expression. Such
CC diseases include immune disorders (e.g. multiple sclerosis, systemic
CC lupus erythematosus and human immuno-deficiency virus (HIV) infections),
CC hyperproliferative disorders (e.g. cancers and Gaucher's disease),
CC cardiovascular diseases (e.g. Scimitar syndrome, Chagas's cardiomyopathy
CC and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC neovascularisation and diabetic retinopathy), neurological disorders
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)
CC and infectious diseases. The polynucleotides and polypeptides are also
CC useful for promoting wound healing, regeneration and/or chemotaxis. The
CC polynucleotides and their complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples. The polypeptides may be used
CC as antigens in the production of antibodies and in assays to identify
CC modulators of protein expression and activity. The anti-TGF-beta receptor
CC antibodies may be used to down regulate expression and activity and as
CC diagnostic agents for detecting the presence of the polypeptides in
CC samples.
XX
SQ Sequence 13 AA;

Query Match 2.9%; Score 13; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLKSRNVV 141
1 VHRDLKSRNVV 13

RESULT 16
ABR71694

ID ABR71694 standard; Protein; 977 AA.

XX ABR71694;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 41874.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL15797.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Disclosure; SEQ ID NO 41874; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABBS7737-ABBS72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 977 AA;

Query Match 2.4%; Score 11; DB 22; Length 977;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLEMLT 201
331 SYGVVLEMLT 341

RESULT 17

ID AAR20470 standard; Protein; 598 AA.

XX AAR20470;

XX 15-APR-1992 (first entry)

XX Soluble ysc alpha.

XX Endoprotease; mutein; carboxypeptidase; ysc alpha; KEX1; KEX2.

XX Saccharomyces cerevisiae.

XX EP467839-A.

XX 22-JAN-1992.

XX 09-UTL-1991; 91EP-0810543.

XX 18-UTL-1990; 90GB-0015825.

XX (CIBA) CIBA GEIGY AG.

XX Heim J, Seebach P, Takabayashi K;

XX WPI; 1992-026641/04.

XX N-PSDB; AAQ20547.

XX In vitro processing of fusion protein - using yeast endo:protease

XX ysc F and carboxypeptidase ysc alpha, for prodn. of biologically

XX active mature protein

XX Example 8; Page 23; 35pp; English.

XX The sequence (SEQ ID NO: 2) is encoded by an insert from plasmid

XX pDPR34/GAPFL-KEX1 and is a sol. form of ysc alpha. The KEX1

XX gene encoding the complete protein (A. Dmochowska et al., Cell 50,

XX 573-584 (1987)) was digested with EcoRI and XhoI to isolate a

XX fragment encoding a C-terminal truncated sol. protein. The gene

XX was expressed in S. cerevisiae. The resulting recombinant sol.

XX enzyme can be used, together with similarly produced sol. yscF

XX for the prepn. of mature proteins by in vitro processing of

XX suitably tailored fusion proteins (see AAR20471, 72).

XX See also AAR20469-72.

XX SQ Sequence 598 AA;

Query Match 2.2%; Score 10; DB 13; Length 598;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDDDD 443
564 SEGDDDDDD 573

DE KEX1 amino acid sequence.

XX KEX1; carboxypeptidase; protease; insulin; ds.

XX Saccharomyces cerevisiae.

XX US4929553-A.

XX 29-MAY-1990.

XX 29-MAY-1987; 87US-005555.

PR 29-MAY-1987; 87US-0055555.
 XX
 PA (CANA) CANADA PATENTS & DEV LTD.
 XX
 PI Bussey H, Dmochowska A, Thomas DY, Dignard D;
 XX
 DR WPI; 1990-192933/25.
 XX N-PSDB; AAQ04768.
 XX
 PT Specific processing of secreted proteins in transformed yeast -
 PT using vector contg. yeast KEX1 gene, encoding protease, and DNA
 PT sequence encoding precursor polypeptide eg. Preproinsulin
 XX
 PS Disclosure; Page ?; ?pp; English.
 XX
 CC Gene product is a protease with carboxypeptidase B activity, useful
 CC in transforming yeast, producing gene products such as prepro
 CC insulin and prepro opio melanocortin.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 729 AA;
 XX
 QY Query Match 2.2%; Score 10; DB 11; Length 729;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 434 SEGDDDDDD 443
 DB 564 SEGDDDDDD 573
 XX
 RESULT 19
 AAM01897
 ID AAM01897 standard; Protein; 764 AA.
 XX
 AC AAM01897;
 XX
 DT 24-NOV-1996 (first entry)
 XX
 DE Nonsense-mediated mRNA decay 2 C-terminal.
 XX
 KM Nonsense-mediated mRNA decay; NMD2; upf1p; inhibition.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO9622301-A1.
 XX
 PD 25-JUL-1996.
 XX
 PF 27-DEC-1995; 95WO-US16930.
 XX
 PR 20-JAN-1995; 95US-0375300.
 XX
 PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
 XX
 PI He F, Jacobson AS;
 XX
 DR WPI; 1996-354469/35.
 DR N-PSDB; AAT131994.
 XX
 PT Isolated nonsense-mediated mRNA decay gene and protein - used to
 PT develop methods for inhibiting the decay pathway for producing
 PT heterologous or endogenous proteins
 PT
 PS Claim 10; Page 54-56; 73pp; English.
 XX
 CC The NMD2 gene (AAT131993) is named after its role in the Nonsense-
 CC Mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upf1p.
 CC A C-terminal fragment of the protein (AAT131994) also binds Upf1p.
 CC when overexpressed in the host cell, the fragment inhibits the
 CC function of Upf1p, thereby inhibiting the nonsense-mediated
 CC mRNA decay pathway.
 CC

SQ Sequence 764 AA;
 XX
 QY Query Match 2.2%; Score 10; DB 17; Length 764;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 437 DDDDDDDGEE 446
 DB 577 DDDDDDDGEE 586
 XX
 RESULT 20
 AAY98056
 ID AAY98056 standard; Protein; 764 AA.
 XX
 AC AAY98056;
 XX
 DT 07-SEP-2000 (first entry)
 XX
 DE Yeast NMD2 carboxy terminal fragment.
 XX
 KM Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;
 KM upf1p; upf1p; gene therapy; beta-thalassemia; cancer;
 KM polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;
 KM haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;
 KM glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;
 KM Cowden disease; Maple syrup urine disease; Wilson disease;
 KM Niemann-Pick disease; Turcot syndrome; McArdle disease;
 KM ornithine transcarbamylase deficiency; antisense therapy; ribozyme.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US6071700-A.
 XX
 PD 06-JUN-2000.
 XX
 PF 21-OCT-1998; 98US-0177431.
 XX
 PR 20-JAN-1995; 95US-0375300.
 PR 21-OCT-1997; 97US-0955472.
 XX
 PA (UYMA-) UNIV MASSACHUSETTS.
 XX
 PI Jacobson AS, He F;
 XX
 DR WPI; 2000-422078/36.
 DR N-PSDB; AAA39451.
 XX
 PT Testing for compounds able to modulate the nonsense mediated mRNA decay
 PT pathway for increasing endogenous protein production by incubating a
 PT candidate compound with a cell having a gene with a nonsense mutation
 PT and a gene without mutation -
 XX
 PS Disclosure; Fig 1; 57pp; English.
 XX
 CC The nonsense-mediated mRNA decay pathway has an important role in the
 CC degradation of mRNA transcripts that contain a stop codon. The products
 CC of the UPF1 and UPF3 genes (proteins Upf1p and Upf3p, respectively) are
 CC essential components of this pathway. Mutations in these genes stabilise
 CC mRNA containing premature stop codons. The present sequence is the
 CC carboxy terminal of the NMD2 gene from Saccharomyces cerevisiae. When
 CC the present sequence is overexpressed, it binds to Upf1p to inhibit its
 CC function. Inhibition of the nonsense-mediated mRNA decay pathway is a
 CC useful means of treating disorders caused by the presence of nonsense
 CC mutations, e.g. breast cancer, polycystic kidney diseases, Niemann-Pick
 CC disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's anaemia,
 CC haemophilia, hypercholesterolaemia, neurofibromatosis, ornithine
 CC transcarbamylase deficiency, retinoblastoma, glycogen storage disease,
 CC McArdle disease, colorectal cancer, Tay-Sachs disease, Cowden disease,
 CC Wilson disease, polycystic kidney disease, Duchenne muscular dystrophy,
 CC adrenocortical carcinoma, Li-Fraumeni syndrome, Maple syrup urine
 CC disease, Turcot syndrome or beta-thalassemia. Antisense mRNA molecules
 CC complementary to the NMD2 gene (AAA39450) may be used for antisense

CC therapy of the above mentioned disorders since the antisense molecules
 CC would inhibit translation of the NMD2 mRNA. Additionally, the NMD2 gene
 CC may be manipulated for ribozyme therapy of the above mentioned
 CC disorders: ribozymes designed to catalytically cleave nonsense-mediated
 CC mRNA decay pathway mRNAs e.g. NMD2 mRNA, can also be used to prevent
 CC translation and therefore expression of these mRNAs.

XX
 SQ Sequence 764 AA;

Query Match 2.2%; Score 10; DB 21; Length 764;
 Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
 DB 577 DDDDDDDGEE 586

RESULT 21

AAW01896
 ID AAW01896 standard; Protein; 1089 AA.

XX
 AC AAW01896;

XX
 DT 24-NOV-1996 (first entry)

XX
 DE Nonsense-mediated mRNA decay 2 protein.

XX
 KM Nonsense-mediated mRNA decay; NMD2; Upf1p; inhibition.

XX
 OS Saccharomyces cerevisiae.

XX
 PN W09622301-A1.

XX
 PD 25-JUL-1996.

XX
 PF 27-DEC-1995; 95MO-US16930.

XX
 PR 20-JAN-1995; 95US-0375300.

XX
 PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.

XX
 PI He F, Jacobson AS;

XX
 DR WPI; 1996-354469/35.

XX
 DR N-PSDB; AATJ1993.

XX
 PT Isolated nonsense-mediated mRNA decay gene and protein - used to
 PT develop methods for inhibiting the decay pathway for producing
 PT heterologous or endogenous proteins

XX
 PS Claim 8; Page 49-52; 73pp; English.

XX
 CC The NMD2 gene (AATJ1993) is named after its role in the Nonsense-
 CC Mediated mRNA Decay pathway. The protein Nmd2p, binds to Upf1p.
 CC A C-terminal fragment of the protein (AATJ1994) also binds to Upf1p.
 CC when overexpressed in the host cell, the fragment inhibits the
 CC function of Upf1p, thereby inhibiting the nonsense-mediated
 CC mRNA decay pathway.

XX
 SQ Sequence 1089 AA;

Query Match 2.2%; Score 10; DB 17; Length 1089;
 Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
 DB 902 DDDDDDDGEE 911

RESULT 22

AAV05835

ID AAV05835 standard; Protein; 1089 AA.

XX
 AC AAV05835;

XX
 DT 02-AUG-1999 (first entry)

XX
 DE Yeast Nmd2p involved in nonsense-mediated mRNA decay pathway.

XX
 KM NMD2 gene; Nmd2p protein; nonsense-mediated mRNA decay pathway;

XX
 KM Yeast; nonsense mutation; therapy; breast cancer; haemophilia;

XX
 KM polycystic kidney disease; Niemann-Pick disease; cystic fibrosis;

XX
 KM adenomatous polyposis coli; Fanconi's anaemia; neurofibromatosis;

XX
 KM hypercholesterolaemia; ornithine transcarbamylase deficiency;

XX
 KM retinoblastoma; glycogen storage disease; McArdle disease; cancer;

XX
 KM Tay-Sachs disease; Cowden disease; Wilson disease;

XX
 KM beta-thalassaemia.

XX
 OS Saccharomyces cerevisiae.

XX
 FH Key location/Qualifiers

XX
 FT Peptide 26..29 /note= "nuclear localisation signal"

XX
 FT Peptide 42..46 /note= "nuclear localisation signal"

XX
 FT Region 843..975 /note= "acidic region"

XX
 FT Protein 326..1089 /note= "C-terminal fragment, inhibits
 XX nonsense-mediated mRNA decay"

XX
 PN W09920797-A1.

XX
 PD 29-APR-1999.

XX
 PF 21-OCT-1998; 98MO-US22365.

XX
 PR 21-OCT-1997; 97US-0955472.

XX
 PA (UYMA-) UNIV MASSACHUSETTS.

XX
 PI He F, Jacobson AS;

XX
 DR WPI; 1999-302756/25.

XX
 DR N-PSDB; AAX25601.

XX
 PT mRNA decay pathway genes and protein useful for treating breast
 XX cancer and cystic fibrosis

XX
 PS Example 3; Fig 1A-C; 116pp; English.

XX
 CC This sequence represents Nmd2p, the protein encoded by the newly
 CC discovered NMD2 gene (see AAX25601) of Saccharomyces cerevisiae.
 CC NMD2 was named after its role in the nonsense-mediated mRNA decay
 CC pathway. The Nmd2p protein binds to Upf1p, another protein in the
 CC decay pathway. A C-terminal fragment of Nmd2p also binds to Upf1p
 CC and, when overexpressed in a host cell, inhibits the function of
 CC Upf1p, thereby inhibiting the nonsense-mediated mRNA decay pathway.
 CC The invention also relates to methods of inhibiting the
 CC nonsense-mediated mRNA decay pathway to stabilise mRNA transcripts
 CC containing a nonsense codon which normally would cause an increased
 CC in the transcript decay rate. Such stabilisation allows an increased
 CC translation from the transcript, and is therefore useful for the
 CC production of a recombinant protein. The invention also relates to
 CC methods of identifying molecules that inhibit the nonsense-mediated
 CC mRNA decay pathway, and the use of such molecules (e.g. the
 CC C-terminal fragment of Nmd2p or an antisense molecule) for

XX
 CC treatment of disorders associated with nonsense mutations, such
 CC as breast cancer, polycystic kidney disease I and II, Niemann-Pick
 CC disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's
 CC anaemia, haemophilia, hypercholesterolaemia, neurofibromatosis,
 CC ornithine transcarbamylase deficiency, retinoblastoma, glycogen
 CC storage disease, McArdle disease, cancer, Tay-Sachs disease, Cowden
 CC disease, Wilson disease or beta-thalassaemia (claimed).

XX SQ Sequence 1089 AA;
Query Match 2.2%; Score 10; DB 20; Length 1089;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDGEE 446
| | | | | | | | | |
| | | | | | | | | |
DB 902 DDDDDDDGEE 911
RESULT 23
AAV98055
ID AAV98055 standard; Protein; 1089 AA.
XX AAV98055;
XX AAV98055;
DT 07-SEP-2000 (first entry)
XX
XX Yeast Nmd2p.
DE
XX Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;
XX upf1p; upf1p; gene therapy; beta-thalassemia; cancer;
XX polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;
XX haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;
XX glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;
XX Cowden disease; Maple syrup urine disease; Wilson disease;
XX Niemann-Pick disease; Turcot syndrome; McArdle disease;
XX ornithine transcarbamylase deficiency; antisense therapy; ribozyme.
XX Saccharomyces cerevisiae.
XX
XX Key Location/Qualifiers
FT Peptide 326..1089
FT /label= Carboxy terminal
FT /note= "Fragment is described in AAV98056"
XX
XX US6071700-A.
XX
XX 06-JUN-2000.
XX
XX 21-OCT-1998; 98US-0177431.
XX
XX 20-JAN-1995; 95US-0375300.
XX 21-OCT-1997; 97US-0955472.
XX
XX (UYMA-) UNIV MASSACHUSETTS.
XX
XX Jacobson AS, He F;
XX WPI, 2000-422078/36.
XX N-PSDB; AAA39450.
XX
XX PT Testing for compounds able to modulate the nonsense mediated mRNA decay
PT pathway for increasing endogenous protein production by incubating a
PT candidate compound with a cell having a gene with a nonsense mutation
PT and a gene without mutation -
XX
XX Disclosure; Fig 1; 57pp; English.
XX
XX The nonsense-mediated mRNA decay pathway has an important role in the
XX degradation of mRNA transcripts that contain a stop codon. The products
XX of the UPF1 and UPF3 genes (proteins Upf1p and Upf3p, respectively) are
XX essential components of this pathway. Mutations in these genes stabilise
XX mRNA containing premature stop codons. The present sequence is the NMD2
XX gene protein from Saccharomyces cerevisiae: Nmd2p. The carboxy terminal
XX of this protein binds to Upf1p to inhibit its function. Inhibition of the
XX nonsense-mediated mRNA decay pathway is a useful means of treating
XX disorders caused by the presence of nonsense mutations, e.g. breast
XX cancer, polycystic kidney diseases, Niemann-Pick disease, adenomatous
XX polyposis coli, cystic fibrosis, Fanconi's anaemia, haemophilia,
XX hypercholesterolaemia, neurofibromatosis, ornithine transcarbamylase

CC deficiency, retinoblastoma, glycogen storage disease, McArdle disease,
CC colorectal cancer, Tay-Sachs disease, Cowden disease, Wilson disease,
CC polycystic kidney disease, Duchenne muscular dystrophy, adrenocortical
CC carcinoma, Li-Fraumeni syndrome, Maple syrup urine disease, Turcot
CC syndrome or beta-thalassemia. Antisense mRNA molecules complementary to
CC the NMD2 gene may be used for antisense therapy of the above mentioned
CC disorders since the antisense molecules would inhibit translation of the
CC NMD2 mRNA. Additionally, the NMD2 gene may be manipulated for ribozyme
CC therapy of the above mentioned disorders: ribozymes designed to
CC catalytically cleave nonsense-mediated mRNA decay pathway mRNAs e.g. NMD2
CC mRNA, can also be used to prevent translation and therefore expression of
CC these mRNAs.
XX
XX SQ Sequence 1089 AA;
Query Match 2.2%; Score 10; DB 21; Length 1089;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDGEE 446
| | | | | | | | | |
| | | | | | | | | |
DB 902 DDDDDDDGEE 911
RESULT 24
AAG21315
ID AAG21315 standard; Protein; 133 AA.
XX AAG21315;
XX AAG21315;
AC AAG21315;
XX
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 23829.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EF1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 08-APR-1999; 99US-0128234.
XX 16-APR-1999; 99US-0128714.
XX 19-APR-1999; 99US-0129845.
XX 21-APR-1999; 99US-0130077.
XX 23-APR-1999; 99US-0130449.
XX 28-APR-1999; 99US-0130510.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 05-MAY-1999; 99US-0132407.
XX 06-MAY-1999; 99US-0132484.
XX 07-MAY-1999; 99US-0132485.
XX 11-MAY-1999; 99US-0132487.
XX 14-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158602.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 434 SEGDDDDDD 442
Db 47 SEGDDDDDD 55

RESULT 25
AAG21314
ID AAG21314 standard; Protein: 156 AA.

XX AAG21314;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 23828.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147203.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159654.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 434 SEGDDDDDD 442
Db 70 SEGDDDDDD 78
RESULT 26
AAAG07429
ID AAAG07429 standard; Protein; 201 AA.
XX
AC AAAG07429;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4579.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

Db 47 SEQDDDDDD 55

RESULT 27

ID AAG07428 standard; Protein; 224 AA.

AC AAG07428;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4578.

XX

KM Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135529.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

```

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      2.0%; Score 9; DB 21; Length 224;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 9; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 434 SEGDDDDDD 442
   |||||
Db 70 SEGDDDDDD 78

```

```

RESULT 28
ABU08115
ID ABU08115 standard; Protein; 224 AA.
XX

```

```

AC ABU08115;
XX
XX 10-MAY-2003 (first entry)
DE Human kinase and phosphatase protein, KPP-16, INCYTE NO. 7510576CD1.
XX
XX Human; enzyme; kinase and phosphatase; KPP; cancer; cirrhosis;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX hepatitis; paroxysmal nocturnal haemoglobinuria; polycythemia vera;
XX psoriasis; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; AIDS; epilepsy;
XX neurological disorder; Alzheimer's disease; Parkinson's disease;
XX autoimmune disorder; inflammatory disorder; allergy; asthma;
XX acquired immunodeficiency syndrome; autoimmune thyroiditis;
XX contact dermatitis; Crohn's disease; diabetes mellitus;
XX glomerulonephritis; Goodpasture's syndrome; gout; Graves' disease;
XX Hashimoto's thyroiditis; irritable bowel syndrome; multiple sclerosis;
XX osteoarthritis; osteoporosis; pancreatitis; Reiter's syndrome;
XX rheumatoid arthritis; Sjogren's syndrome; uveitis; infection.
XX
OS Homo sapiens.
XX
XX WO2003012065-A2.
XX
XX 13-FEB-2003.
XX
XX 01-AUG-2002; 2002WO-US24521.
XX
XX 02-AUG-2001; 2001US-309627P.
XX 07-AUG-2001; 2001US-310933P.
XX 09-AUG-2001; 2001US-311323P.
XX 07-SEP-2001; 2001US-317820P.
XX 14-SEP-2001; 2001US-322264P.
XX 28-SEP-2001; 2001US-326098P.
XX 19-DEC-2001; 2001US-343007P.
XX 15-MAR-2002; 2002US-364494P.
XX 24-APR-2002; 2002US-375539P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughn MR, Yue H, Walla NK, He A, Au-Young JK, Lee SY;
XX Gietzen KJ, Lal PG, Elliott VS, Ison CH, Yang J, Lee EA, Li JX;
XX Emerling BM, Richardson TW, Warren BA, Hatfield AJA, Marquis JP;
XX WPI; 2003-239519/23.
XX N-PSDB; ABX13162.
XX
XX New human kinases and phosphatases and polynucleotides, useful for
XX diagnosing, treating or preventing autoimmune or inflammatory disorders
XX (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
XX cancer or hepatitis -
XX
XX Claim 1; Page 181-182; 1999P; English.
XX
XX The invention relates to an isolated polypeptide, which is a human kinase
XX and phosphatase, KPP (KPP-16). Also included are the
XX encoding polynucleotides KPP NA, recombinant polynucleotide comprising a
XX promoter sequence operably linked to KPP NA, a cell transformed with the
XX recombinant polynucleotide, a transgenic organism comprising the
XX recombinant polynucleotide, an anti-KPP antibody, screening for
XX anti/agonists of KPP, screening for compounds which bind to or alter the
XX activity or expression of KPP, microarray where at least one element is
XX KPP NA, generating an expression profile of a sample containing
XX polynucleotides and an array comprising different nucleotide molecules
XX affixed in distinct physical locations on a solid substrate (where at
XX least one of the nucleotide molecules comprises a first oligonucleotide
XX or polynucleotide sequence specifically hybridizable with at least 30
XX contiguous nucleotides of a target KPP NA). The kinases and phosphatases
XX (KPP) polypeptides, polynucleotides, agonists and antagonists are useful
XX for diagnosing, treating or preventing disorders associated with aberrant
XX expression of KPP, particularly cell proliferative disorders (e.g.
XX arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal
XX nocturnal haemoglobinuria, polycythemia vera, psoriasis, primary

```

CC thrombocytopenia or cancer), developmental disorders (renal tubular
 CC acidosis, anaemia or mental retardation), neurological disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
 CC inflammatory disorders (e.g. AIDS, acquired immunodeficiency
 CC syndrome, allergies, asthma, autoimmune thyroiditis,
 CC contact dermatitis, Crohn's disease, diabetes mellitus,
 CC glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,
 CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
 CC bacterial, fungal, parasitic, protozoan or helminthic infections. They
 CC are also useful in the assessment of the effects of exogenous compounds
 CC on the expression of nucleic acid and aa sequences of proteins
 CC associated with KRP. The polynucleotides encoding KRP are useful for
 CC creating transgenic animals to model human disease. The present
 CC sequence represents a KRP protein of the invention.

SQ Sequence 224 AA;

Query Match 2.0%; Score 9; DB 24; Length 224;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSGGSYVRA 33
 |||||
 Db 17 GSGGSYVRA 25

RESULT 29
 AAR5371
 ID AAR5371 standard; Protein; 502 AA.

AC AAR5371;
 DT 25-MAR-2003 (updated)
 DT 20-JAN-1995 (first entry)

DE Mouse Activin receptor-like kinase 1 (mALK-1).

KW serine threonine kinases; activin receptors; Act-R; superfamily;
 KW transforming growth factor; TGF; diagnostics; detection; therapy;
 KW rheumatoid arthritis; glomerular nephritis; fibrosis.

OS Mus musculus.

PN W09411502-A2.

PD 26-MAY-1994.

PF 17-NOV-1993; 93WO-GB02367.

XX 17-NOV-1992; 92GB-0024057.

PR 08-MAR-1993; 93GB-0004677.

PR 08-MAR-1993; 93GB-0004680.

PR 28-MAY-1993; 93GB-0011047.

PR 02-JUL-1993; 93GB-0013763.

PR 03-AUG-1993; 93GB-0016099.

PR 15-OCT-1993; 93GB-0021344.

XX (LUDW-) LUDWIG INST CANCER RES.

PA Dijke P, Franzen P, Heldin C, Miyazono K, Yamashita H;

DR WPI; 1994-183503/22.

DR N-PSDB; AAO66671.

XX Actin receptor-like kinase(s) with serine/threonine kinase
 PT domains - have activin/TGF beta type I receptor function and can
 PT be used in diagnosis or therapy of rheumatoid arthritis,
 PT glomerular nephritis, fibrosis, etc.
 XX Claim 3; Page 60-61; 97pp; English.

CC The inventors have identified a new family of receptor kinases
 CC called activin receptor-like kinases (ALK). Their discovery was
 CC based on the realisation that receptor serine/threonine kinases
 CC form a new receptor family, which may include the type II receptors
 CC for other proteins in the transforming growth factor (TGF) beta
 CC superfamily. The activin receptor type II sequences from mouse and
 CC the dafI gene product of C.elegans have high sequence similarity
 CC and were used to design degenerate primers to clone related cDNA's
 CC (see AAO66643-49). Six distinct putative receptor serine/threonine
 CC kinases were identified, called ALK (human ALK proteins are shown
 CC in AAR5366-70, mouse ALK are shown in AAR5371-74). Products of the
 CC invention can be used in therapy, eg. to modulate conditions
 CC associated with activin or TGF beta activity. These conditions
 CC include fibrosis, eg. liver cirrhosis and pulmonary fibrosis, cancer,
 CC rheumatoid arthritis and glomerulonephritis.
 CC (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 502 AA;

Query Match 2.0%; Score 9; DB 15; Length 502;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 HRDLKSRNV 139
 |||||
 Db 327 HRDLKSRNV 335

RESULT 30
 AAY3304
 ID AAY3304 standard; Protein; 502 AA.

AC AAY3304;
 DT 26-NOV-1999 (first entry)

DE Human mALK-1 clone AM6 protein.

KW mALK-1; activin receptor-like kinase; therapy; activin; TGF-beta; cancer;
 KW fibrosis; liver cirrhosis; pulmonary fibrosis; glomerulonephritis; mouse;
 KW rheumatoid arthritis; detection; diagnosis; drug screening.

OS Mus sp.

PN W0946386-A1.

PD 16-SEP-1999.

PF 12-MAR-1999; 99WO-US05735.

XX 13-MAR-1998; 98US-0039177.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Miyazono K, Imamura T, Ten Dijke P;

DR WPI; 1999-551413/46.

DR N-PSDB; AA209846.

XX New isolated activin receptor-like kinases, used to develop products

PT for treating e.g. fibrosis, cancer, rheumatoid arthritis and

PT glomerulonephritis -

XX Disclosure; Page 86-88; 110pp; English.

XX This invention describes novel human and murine activin receptor-like

CC kinases (ALK's). The novel ALK products can be used in therapy, e.g. to

CC modulate conditions associated with activin or TGF-beta activity, such as

CC fibrosis, e.g. liver cirrhosis and pulmonary fibrosis, cancer, rheumatoid

CC arthritis and glomerulonephritis. The products can also be used for

CC detection, diagnosis and drug screening. This sequence represents the

CC murine mALK-1 protein isolated from clone AM6.

SQ Sequence 502 AA;
 Query Match 2.0%; Score 9; DB 20; Length 502;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 131 HRDLKSRNV 139
 Db 327 HRDLKSRNV 338
 RESULT 31
 AAR41922
 ID AAR41922 standard; Protein; 505 AA.
 XX
 AC AAR41922;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-APR-1994 (first entry)
 XX
 DE MISR3.
 XX
 KM MIS; Mullerian Inhibitory Substance; receptor;
 KM transforming growth factor; inhibin; BMP;
 KM membrane serine/threonine kinase receptor;
 KM bone morphogenesis protein.
 XX
 OS Rattus rattus.
 XX
 PN WO9319177-A1.
 XX
 PD 30-SEP-1993.
 XX
 PF 15-MAR-1993; 93WO-US02387.
 XX
 PR 18-MAR-1992; 92US-0853396.
 PR 11-MAR-1993; 93US-0029673.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Donahoe PK, Gustafson M, He WW;
 XX
 DR WPI; 1993-320743/40.
 DR N-PSDB; AAR41922.
 XX
 PT New receptors of the transforming growth factor-beta receptor
 PT family - comprising Mullerian Inhibitory Substance Receptors and
 PT inhibin receptors
 XX
 PS Claim 51; Fig 3; 59pp; English.
 XX
 CC Mistr1 (AAQ49763) is believed to encode an isoform of the rat
 CC MIS receptor. Mistr2A/mistr2B (AAQ49764), mistr3 (AAQ49765) and mistr4
 CC (AAQ49766) are believed to encode monomeric isoforms of the rat
 CC inhibin receptor and/or BMP receptor.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 SQ Sequence 505 AA;
 Query Match 2.0%; Score 9; DB 14; Length 505;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 131 HRDLKSRNV 139
 Db 330 HRDLKSRNV 338
 RESULT 32
 AAW03761
 ID AAW03761 standard; Protein; 505 AA.
 XX
 AC AAW03761;

XX
 DT 25-MAR-2003 (updated)
 DT 30-OCT-1996 (first entry)
 XX
 DE Mullerian inhibiting substance receptor MISR3.
 XX
 KM Mullerian inhibiting substance receptor; MISR; TGF-beta receptor;
 KM transforming growth factor beta type I receptor; gene therapy;
 KM wound healing; tumour treatment; rat inhibin.
 XX
 OS Rattus sp.
 XX
 PN US538892-A.
 XX
 PD 23-JUL-1996.
 XX
 PF 04-NOV-1993; 93US-0149105.
 XX
 PR 04-NOV-1993; 93US-0149105.
 PR 18-MAR-1992; 92US-0853396.
 PR 11-MAR-1993; 93US-0029673.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (UYDU-) UNIV DUKE.
 XX
 PI Donahoe PK, Gustafson M, He W, Wang X;
 XX
 DR WPI; 1996-353830/35.
 DR N-PSDB; AAT36071.
 XX
 PT New isolated TGF-beta type I receptor DNA - used to develop prods
 PT for diagnosis and therapy, e.g. for treating tumours or promoting
 PT wound healing
 XX
 PS Disclosure; Columns 45-48; 44pp; English.
 XX
 CC Degenerate PCR primers were designed based on two highly conserved
 CC regions within the cDNA encoding a murine activin receptor, human
 CC and porcine TGF-beta type II receptor and the daf-1 receptor of
 CC C.elegans. The primers (see AAT36072 and AAT36073) were used for
 CC amplifying clones present in a 14.5 day foetal rat urogenital ridge
 CC cDNA COS cell expression library. Four clones encoding portions of
 CC four novel polypeptides (all putative serine/threonine kinases)
 CC were obtained and inserts from these clones were used as probes to
 CC isolate full-length cDNA sequences for each of the four TGF-beta
 CC type I receptors. Each putative receptor comprises a hydrophobic
 CC signal peptide of 19-23 residues, an extracellular, cysteine-rich,
 CC hydrophilic, ligand-binding domain of 100-150 residues, a
 CC hydrophobic single transmembrane domain of 23-25 residues, an
 CC intracellular serine/threonine kinase domain of approximately 300
 CC residues and a short serine/threonine-rich tail. The present sequence
 CC is that of MISR3 which is believed to be a monomeric isoform of the
 CC rat inhibin receptor and/or BMP receptor.
 CC Isolated DNA which encodes the MISR4 sequence or which is able to
 CC hybridize to such DNA under stringent conditions is claimed.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 XX
 SQ Sequence 505 AA;
 Query Match 2.0%; Score 9; DB 17; Length 505;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 131 HRDLKSRNV 139
 Db 330 HRDLKSRNV 338

RESULT 33

ID ABP30294 standard; Protein; 648 AA.

AC ABP30294;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 9764.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettein H;

XX WPI; 2002-352536/38.

XX N-PSDB; ABN70925.

XX New Streptococcus protein for the treatment or prevention of infection

XX or disease caused by Streptococcus bacteria, such as meningitis, and

XX for detecting a compound that binds to the protein -

XX Claim 1; Page 4103; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX the specification. The proteins have antibacterial and antiinflammatory

XX activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX antibodies that bind (I) are used in the manufacture of medicaments for

XX the treatment or prevention of infection or disease caused by

XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX Nucleic acids encoding (I) are used to detect Streptococcus in a

XX biological sample. (I) is used to determine whether a compound binds to

XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX used as a vaccine or diagnostic composition. The disease caused by

XX Streptococcus that is prevented or treated may be meningitis. Nucleic

XX acid encoding (I) may be used to recombinantly produce (I) and may be

XX used in gene therapy. Antibodies to (I) are used for affinity

XX chromatography, immunoassays, and distinguishing/identifying

XX Streptococcus proteins.

XX Sequence 648 AA;

XX Query Match 2.0%; Score 9; DB 23; Length 648;

XX Best Local Similarity 100.0%; Pred. No. 13;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 51 EKEAEILSV 59

XX 303 EKEAEILSV 311

XX RESULT 34

ABP27599

ID ABP27599 standard; Protein; 652 AA.

AC ABP27599;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4374.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettein H;

XX WPI; 2002-352536/38.

XX N-PSDB; ABN68230.

XX New Streptococcus protein for the treatment or prevention of infection

XX or disease caused by Streptococcus bacteria, such as meningitis, and

XX for detecting a compound that binds to the protein -

XX Claim 1; Page 3598; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX the specification. The proteins have antibacterial and antiinflammatory

XX activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX antibodies that bind (I) are used in the manufacture of medicaments for

XX the treatment or prevention of infection or disease caused by

XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX Nucleic acids encoding (I) are used to detect Streptococcus in a

XX biological sample. (I) is used to determine whether a compound binds to

XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX used as a vaccine or diagnostic composition. The disease caused by

XX Streptococcus that is prevented or treated may be meningitis. Nucleic

XX acid encoding (I) may be used to recombinantly produce (I) and may be

XX used in gene therapy. Antibodies to (I) are used for affinity

XX chromatography, immunoassays, and distinguishing/identifying

XX Streptococcus proteins.

XX Sequence 652 AA;

XX Query Match 2.0%; Score 9; DB 23; Length 652;

XX Best Local Similarity 100.0%; Pred. No. 13;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 51 EKEAEILSV 59

XX 307 EKEAEILSV 315

XX RESULT 35

XX ABP27600 standard; Protein; 652 AA.

AC	ABP27600;
XX	
DT	02-JUL-2002 (first entry)
XX	
DE	Streptococcus polypeptide SEQ ID NO 4376.
XX	
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM	group A streptococcus; Streptococcus pyogenes; antibacterial;
KX	antiinflammatory; infection; vaccine; meningitis; gene therapy.
OS	Streptococcus pyogenes.
XX	
PN	MO200234771-A2.
PD	
FD	02-MAY-2002.
XX	
PF	29-OCT-2001; 2001WO-GB04789.
XX	
PR	27-OCT-2000; 2000GB-0026333.
PR	24-NOV-2000; 2000GB-0028727.
PR	07-MAR-2001; 2001GB-0005640.
XX	
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	
P1	Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
P1	Tetcelin H,
XX	
DR	WPI; 2002-352536/38.
DR	N-PSDB; ABN68231.
XX	
PT	New Streptococcus protein for the treatment or prevention of infection
PT	or disease caused by Streptococcus bacteria, such as meningitis, and
PT	for detecting a compound that binds to the protein -
XX	
PS	Claim 1; Page 3598; 4525pp; English.
XX	
CC	The invention relates to a protein (ABP25413-ABP30895) from group B
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC	the specification. The proteins have antibacterial and antiinflammatory
CC	activity. (II) nucleic acids encoding (I), ABN6004-ABN72526 and
CC	antibodies that bind (I) are used in the manufacture of medicaments for
CC	the treatment or prevention of infection or disease caused by
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a
CC	biological sample. (I) is used to determine whether a compound binds to
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC	used as a vaccine or diagnostic composition. The disease caused by
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic
CC	acid encoding (I) may be used to recombinantly produce (I) and may be
CC	used in gene therapy. Antibodies to (I) are used for affinity
CC	chromatography, immunoassays, and distinguishing/identifying
CC	Streptococcus proteins.
SQ	Sequence 652 AA:
OY	Query Match 2.0%; Score 9; DB 23; Length 652;
DB	Best Local Similarity 100.0%; Pred. No. 13;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
51 EXEAEILSV 59	
307 EXEAEILSV 315	
RESULT 36	
ID AAE16259 standard; Protein; 656 AA.	
AAE16259	
AAE16259;	
26-MAR-2002 (first entry)	

DE	XX	Human kinase PKIN-5 protein.
XX	XX	Human, kinase, PKIN-5, cancer, leukaemia, adenocarcinoma, osteoarthritis
KW	KW	immune disorder, atherosclerosis, Crohn's disease, Hodgkin's disease;
KW	KW	Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
KW	KW	allergy; asthma; adult respiratory distress syndrome; multiple sclerosis
KW	KW	autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;
KW	KW	Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
KW	KW	rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
KW	KW	hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
KW	KW	cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
KW	KW	congestive heart failure; ischaemic heart disease; lung tumour; gout;
KW	KW	fatty liver; Niemann-Pick's disease; gene therapy.
XX	OS	Homo sapiens.
XX	XX	
FT	PH	Key
FT	FT	Location/Qualifiers
FT	FT	Domain
FT	FT	/label= Protein_kinase_domain
FT	FT	14..197
FT	FT	/label= Protein_kinase_domain
FT	FT	14..257
FT	FT	/note= "Eukaryotic protein kinase domain"
FT	FT	Domain
FT	FT	14..252
FT	FT	/label= Protein_kinase_domain
FT	FT	14..253
FT	FT	/label= Protein_kinase_domain
XX	PN	MO200196547-A2.
XX	XX	
PD	XX	20-DEC-2001.
XX	PF	14-JUN-2001; 2001MO-US19444.
XX	XX	
PR	PR	15-JUN-2000; 2000US-212073P.
PR	PR	23-JUN-2000; 2000US-213467P.
PR	PR	30-JUN-2000; 2000US-215651P.
PR	PR	07-JUL-2000; 2000US-216605P.
PR	PR	13-JUL-2000; 2000US-218372P.
PR	PR	25-AUG-2000; 2000US-228056P.
XX	XX	
PA	XX	(INCY-) INCYTE GENOMICS INC.
XX	XX	
P1	P1	Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
P1	P1	Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR;
P1	P1	Ramkumar J, Griffin JA, Kearney L, Buford N, Nguyen DB, Tang YT;
P1	P1	Baughn MB, He A, Thornton M, Hafalla A, Patterson C, Gururajan R;
P1	P1	Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
P1	P1	Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX	XX	
DR	DR	WPI; 2002-090207/12.
DR	DR	PSDB; AAD26452.
XX	XX	
PT	PT	New polypeptides, useful for diagnosing, treating or preventing
PT	PT	disorders of growth and development, cardiovascular and lipid, and
PT	PT	diseases such as cancer, comprise human kinase polypeptides -
XX	XX	
PS	PS	Claim 1; Page 143-144; 197pp; English.
XX	XX	
CC	CC	The invention relates to human kinase PKIN proteins and their
CC	CC	corresponding cDNAs. A composition containing PKIN agonist is useful for
CC	CC	treating a disease or condition associated with decreased expression of
CC	CC	PKIN and a composition comprising PKIN antagonist is useful for treating
CC	CC	a disease or condition associated with overexpression of PKIN. The
CC	CC	disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,
CC	CC	myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC	CC	(Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC	CC	atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC	CC	autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC	CC	mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC	CC	osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC	CC	rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC	CC	

CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
 CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypercholesterolaemia, obesity). PKIN DNA is useful for assessing
 CC toxicity of a test compound and in gene therapy. The present sequence
 CC is human PKIN-5 protein.

XX
 XX
 SQ Sequence 656 AA;

Query Match 2.0%; Score 9; DB 23; Length 656;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSYVRA 33
 |||||
 Db 17 GSFGSYVRA 25

RESULT 37
 AAU03554
 ID AAU03554 standard; Protein; 909 AA.
 XX
 AC AAU03554;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human protein kinase #54.
 XX
 XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 XX reproductive disorder.
 OS Homo sapiens.
 XX
 PN WO200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32085.
 XX
 PR 24-NOV-1999; 99US-0167482.
 XX
 PA (SUGB-) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 DR MPI; 2001-343950/36.
 XX
 PT N-PSDB; AAS06754.
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 XX Claim 7; Figure 2; 433p; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).

CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 XX
 SQ Sequence 909 AA;

Query Match 2.0%; Score 9; DB 22; Length 909;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSYVRA 33
 |||||
 Db 403 GSFGSYVRA 411

RESULT 38
 ABU1054
 ID ABU1054 standard; Protein; 911 AA.
 XX
 AC ABU1054;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human protein NOV19b.
 XX
 XX Human; NOVX; adrenoleukodystrophy; haemophilia; stroke; VHL;
 KW congenital adrenal hyperplasia; haemophilia; hypercoagulation;
 KW idiopathic thrombocytopenic purpura; autoimmune disease; allergy;
 KW immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;
 KW Alzheimer's disease; tubular sclerosis; Parkinson's disease; epilepsy;
 KW Huntington's disease; cerebral palsy; Leach-Ryan syndrome; pain;
 KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;
 KW behavioural disorder; addiction; neuroprotection; diabetes; ARDS;
 KW renal artery stenosis; interstitial nephritis; glomerulonephritis;
 KW polycystic kidney disease; systemic lupus erythematosus; IGA;
 KW renal tubular acidosis; immunoglobulin A nephropathy; hypercalcaemia;
 KW cirrhosis; transplantation; ascites; emphysema; scleroderma; GVHD;
 KW adult respiratory distress syndrome; graft versus host disease;
 KW lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;
 KW anaemia; cancer; trauma; regeneration; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200281629-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 03-APR-2002; 2002WO-US10522.
 XX
 PR 03-APR-2001; 2001US-281086P.
 PR 03-APR-2001; 2001US-281136P.
 PR 05-APR-2001; 2001US-281863P.
 PR 05-APR-2001; 2001US-281906P.
 PR 10-APR-2001; 2001US-282934P.
 PR 12-APR-2001; 2001US-283512P.
 PR 19-APR-2001; 2001US-285325P.
 PR 23-APR-2001; 2001US-285890P.
 PR 24-APR-2001; 2001US-286068P.
 PR 25-APR-2001; 2001US-286282P.
 PR 12-MAY-2001; 2001US-281134P.
 PR 15-MAY-2001; 2001US-282020P.
 PR 17-MAY-2001; 2001US-291725P.
 PR 31-MAY-2001; 2001US-294771P.
 PR 08-JUN-2001; 2001US-296965P.
 PR 18-JUN-2001; 2001US-299128P.
 PR 12-UTL-2001; 2001US-305063P.
 PR 14-NOV-2001; 2001US-332780P.
 PR 04-JAN-2002; 2002US-345221P.
 XX
 PA (CURA-) CURAGEN CORP.

PI Szytek KA, Li L, Edinger SR, Ellerman K, Stone DJ, Malyankar UM,
PI Shinkets RA, Guo X, Anderson DW, Paturajan M, Bergts C,
PI Geilach V, Taupier RJ, Pena CE, Padigar M, Liu Y, Burgess CE,
PI Miller CE, Gusev VY, Kekuda R, Gorman L, Zernusen BD,
PI Baumgartner JC, Tcherner VT, Vernet CAM, Smithson G, Heyes MP,
PI Shenoy SG, Liu X, Gangoli EA;
XX
XX WPI: 2003-046863/04.
DR N-PSDB; AEX17532.
DR
XX
PT New polypeptides, designated NOVX polypeptides, useful for treating
PT hemophilia, idiopathic thrombocytopenic purpura, autoimmune disease,
PT allergies, transplantation, Alzheimer's disease and stroke -
XX
XX Claim 1; Page 146; 320pp; English.

Seq	Sequence	911 AA;
	Query Match	2.0%;
	Best Local Similarity	100.0%;
	Pred. No.	17;
	Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	25 GSFGSVYRA	33
DB	282 GSFGSVYRA	290

RESULT	39
ABU11053	
ID	ABU11053 standard; Protein, 915 AA.
XX	
AC	ABU11053;
XX	
DT	05-FEB-2003 (first entry)
XX	
DE	Human protein NOV19a.
XX	
KW	Human; NOX; adrenoleukodystrophy; haemophilia; stroke; VHL;
KW	congenital adrenal hyperplasia; haemophilia; hypercoagulation;
KW	idiopathic thrombocytopenic purpura; autoimmune disease; allergy;
KW	immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;

KW Alzheimer's disease; tubular sclerosis; Parkinson's disease; epilepsy;
 KM Huntington's disease; cerebral palsy; Lesch-Nyhan syndrome; pain;
 KM multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;
 KM behavioural disorder; addiction; neuroprotection; diabetes; ARDS;
 KM renal artery stenosis; interstitial nephritis; glomerulonephritis;
 KM polycystic kidney disease; systemic lupus erythematosus; IGA;
 KM renal tubular acidosis; immunoglobulin A nephropathy; hypercalcemia;
 KM cirrhosis; transplantation; asthma; emphysema; scleroderma; GVHD;
 KM adult respiratory distress syndrome; graft versus host disease;
 KM lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;
 KW anaemia; cancer; trauma; regeneration; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200281629-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 03-APR-2002; 2002WO-US10522.
 XX
 PR 03-APR-2001; 2001US-281086P.
 PR 03-APR-2001; 2001US-281136P.
 PR 05-APR-2001; 2001US-281863P.
 PR 05-APR-2001; 2001US-281906P.
 PR 10-APR-2001; 2001US-282934P.
 PR 12-APR-2001; 2001US-283512P.
 PR 19-APR-2001; 2001US-285325P.
 PR 23-APR-2001; 2001US-285890P.
 PR 24-APR-2001; 2001US-286068P.
 PR 25-APR-2001; 2001US-286292P.
 PR 12-MAY-2001; 2001US-291134P.
 PR 15-MAY-2001; 2001US-282020P.
 PR 17-MAY-2001; 2001US-291725P.
 PR 31-MAY-2001; 2001US-294771P.
 PR 08-JUN-2001; 2001US-296965P.
 PR 18-JUN-2001; 2001US-299128P.
 PR 12-JUL-2001; 2001US-305063P.
 PR 14-NOV-2001; 2001US-332780P.
 PR 04-JAN-2002; 2002US-345221P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Spytek KA, Li L, Edinger SR, Ellerman K, Stone DJ, Malyankar UM;
 PI Shimkets RA, Guo X, Anderson DW, Patutarajan M, Berghe C;
 PI Gerlach V, Taupier RJ, Pena CE, Padigaru M, Liu Y, Burgess CE;
 PI Miller CE, Gusev VV, Kekuda R, Gorman L, Zernusen BD;
 PI Baumgartner JC, Tchierne VT, Vernet CM, Smithson G, Heyes MP;
 PI Shenoy SG, Liu X, Gangoli EA;
 PI
 XX WPI; 2003-046863/04.
 DR N-PSDB; ABX17531.
 XX
 PT New polypeptides, designated NOVX polypeptides, useful for treating
 PT hemophilia, idiopathic thrombocytopenic purpura, autoimmune disease,
 PT allergies, transplantation, Alzheimer's disease and stroke -
 PT
 XX
 PS Claim 1, Page 144-145; 320pp; English.
 XX
 CC The invention relates to an isolated NOVX polypeptide selected from
 CC NOVX-17527 polypeptides, a mature form of NOVX, a variant of NOVX
 CC or a fragment of NOVX. Also included are determining the presence or
 CC amount of NOVX in a sample (by using an antibody that immunospecifically
 CC bind to the polypeptide), determining the presence of or predisposition
 CC to disease associated with altered levels of NOVX in a first mammalian
 CC subject, identifying a potential therapeutic agent for use in the
 CC treatment of pathology related to aberrant expression of physiological
 CC interactions of NOVX, screening for a modulator of activity of or latency
 CC or predisposition to a pathology associated with NOVX, the nucleic
 CC acid encoding NOVX, vectors and host cells. NOVX is useful for
 CC identifying an agent (a cellular receptor or downstream effector) that
 CC binds to NOVX. NOVX and NOVX nucleic acids are useful for treating or
 CC preventing NOVX-associated disorders in humans, and in the
 CC manufacture of a medicament for treating a NOVX related disease

CC human disease e.g. adrenoleukodystrophy, congenital adrenal
 CC hyperplasia, haemophilia, hypercoagulation, idiopathic thrombocytopenic
 CC purpura, autoimmune disease, allergies, immunodeficiencies,
 CC transplantation, von Hippel-Lindau (VHL) syndrome, Alzheimer's disease,
 CC stroke, tubular sclerosis, Parkinson's disease, Huntington's disease,
 CC cerebral palsy, epilepsy, Leisch-Nyhan syndrome, multiple sclerosis,
 CC ataxia-cerebellar ataxia, leukodystrophies, behavioural disorders,
 CC addition, anxiety, pain, neuroprotection, diabetes, renal artery
 CC stenosis, interstitial nephritis, glomerulonephritis, polycystic kidney
 CC disease, systemic lupus erythematosus, renal tubular acidosis,
 CC immunoglobulin (Ig) A nephropathy, hypercalcaemia, cirrhosis,
 CC transplantation, asthma, emphysema, scleroderma, adult respiratory
 CC distress syndrome (ARDS), graft versus host disease (GVHD), lymphedema,
 CC fertility, pancreatitis, obesity, haemophilia, ulcers, anaemia, cancer,
 CC trauma, regeneration, and viral, bacterial or parasitic infections.
 CC The present sequence represents a NOVX protein.

XX Sequence 915 AA;

Query Match 2.0%; Score 9; DB 24; Length 915;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSEFGSVYRA 33
 |||||
 Db 282 GSEFGSVYRA 290

RESULT 40
 AARS5694
 ID AARS5694 standard; Protein; 2391 AA.

XX AARS5694;

DT 25-MAR-2003 (updated)
 DT 06-DEC-1994 (first entry)

XX Carbamoyl-phosphate-synthetase II.

KW Carbmoyl-phosphate-synthetase II; CPSII; pscCPSII gene;
 KW malaria.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT 1..690 /note="glutamine-amidotransferase domain"

FT 1..270 /note="structural subdomain"

FT 271..482 /note="insert sequence"

FT 483..690 /note="insert sequence"

FT /note="glutamine subdomain"

FT 691..2391 /note="carbamoyl-phosphate-synthetase domain"

FT 691..1254 /note="ATP binding subdomain CPsa"

FT 1255..1857 /note="insert sequence"

FT 1858..2391 /note="ATP binding subdomain CPsb"

XX Domain

XX MO9412643-A1.

XX 09-JUN-1994.

XX 02-DEC-1993; 93WO-AU00617.

XX 03-DEC-1992; 92AU-0006206.

XX 16-DEC-1992; 92AU-0006380.

XX (UNIX) UNISEARCH LTD.

PI Flores MV, Oesullivan WJ, Stewart TS;
 XX MPI, 1994-2002/1/24.
 DR N-PSDB; AAQ62924.
 XX Nucleic acid encoding carbamoyl phosphate synthetase II -
 PT isolated from Plasmodium falciparum, used to develop prods. for
 PT the treatment of malaria.
 XX Disclosure; Page 6-16; 31pp; English.

PS The CDNA sequence encoding the carbamoyl-phosphate-transferase II
 CC (CPSII) of Plasmodium falciparum was determined. The CDNA encodes
 CC a protein that includes 2 insert sequences not found in other CPSII
 CC proteins. The first separates the putative structural subdomain and
 CC the glutamine subdomain of the glutamine-amidotransferase subunit
 CC of CPSII, while the second separates 2 ATP binding subdomains of the
 CC CPSII subunit, CPsa and CPsb.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 2391 AA;

Query Match 2.0%; Score 9; DB 15; Length 2391;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445
 |||||
 Db 1786 DDDDDDDGE 1794

RESULT 41
 ABU08105
 ID ABU08105 standard; Protein; 2527 AA.

XX ABU08105;

DT 10-MAY-2003 (first entry)

XX Human kinase and phosphatase protein, KPP-6, INCYTE No. 71059650CD1.

XX Human; enzyme; kinase and phosphatase; KPP; cancer; cirrhosis;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 KW hepatitis; paroxysmal nocturnal haemoglobinuria; polycythemia vera;
 KW psoriasis; primary thrombocytopaenia; developmental disorder;
 KW renal tubular acidosis; anaemia; mental retardation; AIDS; epilepsy;
 KW neurological disorder; Alzheimer's disease; Parkinson's disease;
 KW autoimmune disorder; inflammatory disorder; allergy; asthma;
 KW acquired immunodeficiency syndrome; autoimmune thyroiditis;
 KW contact dermatitis; Crohn's disease; diabetes mellitus;
 KW glomerulonephritis; Goodpasture's syndrome; gout; Graves' disease;
 KW Hashimoto's thyroiditis; irritable bowel syndrome; multiple sclerosis;
 KW osteoarthritis; osteoporosis; pancreatitis; Reiter's syndrome;
 KW rheumatoid arthritis; Sjogren's syndrome; uveitis; infection.

XX Homo sapiens.

XX MO2003012065-A2.

XX 13-FEB-2003.

XX 01-AUG-2002; 2002WO-US24521.

XX 02-AUG-2001; 2001US-309627P.

XX 07-AUG-2001; 2001US-310933P.

XX 09-AUG-2001; 2001US-311323P.

XX 07-SEP-2001; 2001US-317820P.

XX 14-SEP-2001; 2001US-322284P.

XX 28-SEP-2001; 2001US-326098P.

XX 19-DEC-2001; 2001US-343007P.

XX 15-MAR-2002; 2002US-364494P.

XX 24-APR-2002; 2002US-375539P.

PT soybean, by preparing lunasin gene constructs using Pichia expression
PT vectors, optimizing lunasin expression parameters and purifying lunasin
XX
XX
PS Claim 7; Page 62; 67pp; English.
XX
CC The invention provides a method of producing lunasin, a cancer preventive
CC peptide from soybean, or its active variant by recombinant DNA technology
CC in large quantities. The method comprises preparing lunasin gene
CC constructs or its variants using protein expression vectors, optimizing
CC lunasin expression parameters for large-scale production and performing
CC a series of isolation and purification to obtain large quantities of
CC biologically active recombinant lunasin peptide. The lunasin peptide
CC exhibits inhibitory effects against malignant transformation of cells
CC induced by chemical carcinogens and viral oncogenes. Sequences
CC AAB62614-24 represent soybean lunasin variants.
XX
SQ Sequence 20 AA;

Query Match 1.8%; Score 8; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 13 GDDDDDD 20

RESULT 44
AAB83386
ID AAB83386 standard; peptide; 21 AA.
XX
AC AAB83386;
XX
DT 26-MAR-2002 (first entry)
XX
DE Lunasin fragment #7.
XX
KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
XX prostate; liver; kidney; therapy; soybean.
XX
OS Glycine max.
XX
PN MO200172784-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001MO-US09453.
XX
PR 24-MAR-2000; 2000US-0534705.
XX
PA (FIG-) PILGEN BIOSCIENCES INC.
XX
PI Galvez AF;
XX
DR WPI; 2001-648438/74.
XX
PT Treatment or prevention of cancer by administering a lunasin peptide
XX which binds to non-acetylated histones H3 and H4 and prevents their
XX acetylation in hypoacetylated repressed chromatin -
PS Claim 7; Page 42; 49pp; English.
XX
CC This sequence represents a lunasin peptide used in the method of the
CC invention. The method is for the treatment or prevention of cancer,
CC comprising administration of a lunasin peptide or its active fragment or
CC analog. The invention is used to treat or prevent cancer, particularly
CC in a human. The cancers that can be treated and prevented include
CC those of the colon, upper gastrointestinal tract, breast, prostate,
CC liver, kidney or any other internal organs or tissues.
XX
SQ Sequence 21 AA;

Query Match 1.8%; Score 8; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 13 GDDDDDD 20

RESULT 45
AAB83393
ID AAB83393 standard; peptide; 21 AA.
XX
AC AAB83393;
XX
DT 26-MAR-2002 (first entry)
XX
DE Lunasin related peptide #2.
XX
KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
XX prostate; liver; kidney; therapy; soybean.
XX
OS Glycine max.
XX
PN MO200172784-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001MO-US09453.
XX
PR 24-MAR-2000; 2000US-0534705.
XX
PA (FIG-) PILGEN BIOSCIENCES INC.
XX
PI Galvez AF;
XX
DR WPI; 2001-648438/74.
XX
PT Treatment or prevention of cancer by administering a lunasin peptide
XX which binds to non-acetylated histones H3 and H4 and prevents their
XX acetylation in hypoacetylated repressed chromatin -
XX
PS Disclosure; Page 46; 49pp; English.
XX
CC This sequence represents a lunasin related peptide used in the method of
CC the invention. The method is for the treatment or prevention of cancer,
CC comprising administration of a lunasin peptide or its active fragment or
CC analog. The invention is used to treat or prevent cancer, particularly
CC in a human. The cancers that can be treated and prevented include
CC those of the colon, upper gastrointestinal tract, breast, prostate,
CC liver, kidney or any other internal organs or tissues.
XX
SQ Sequence 21 AA;

Query Match 1.8%; Score 8; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 12 GDDDDDD 19

RESULT 46
AAB62620
ID AAB62620 standard; peptide; 21 AA.
XX
AC AAB62620;
XX
DT 23-JUL-2001 (first entry)
XX
DE Soybean lunasin peptide variant (residues 22-42).
XX

KM Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.
 XX Glycine max.
 OS
 XX
 PN WO200134808-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 12-NOV-2000; 2000WO-US31211.
 XX
 PR 12-NOV-1999; 99US-0165334.
 XX
 PA (FILG-) FILGEN BIOSCIENCES INC.
 XX
 PI Galvez AF;
 PT WPI; 2001-343605/36.
 XX
 PT Large scale production of lunasin, a cancer preventive peptide from
 PT soybean, by preparing lunasin gene constructs using Pichia expression
 PT vectors, optimizing lunasin expression parameters and purifying lunasin
 PT
 XX
 PS Claim 7; Page 62; 67pp; English.
 XX
 CC The invention provides a method of producing lunasin, a cancer preventive
 CC peptide from soybean, or its active variant by recombinant DNA technology
 CC in large quantities. The method comprises preparing lunasin gene
 CC constructs or its variants using protein expression vectors, optimizing
 CC lunasin expression parameters for large-scale production and performing
 CC a series of isolation and purification to obtain large quantities of
 CC biologically active recombinant lunasin peptide. The lunasin peptide
 CC exhibits inhibitory effects against malignant transformation of cells
 CC induced by chemical carcinogens and viral oncogenes. Sequences
 CC AAB62614-24 represent soybean lunasin variants.
 CC
 XX
 SQ Sequence 21 AA;
 XX
 Query Match 1.8%; Score 8; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 436 GDDDDDD 443
 DB 13 GDDDDDD 20
 XX
 RESULT 47
 AABG50623
 ID AABG50623 standard; Peptide; 22 AA.
 XX
 AC AABG50623;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID No 29271.
 XX
 KM Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KM hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 XX

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 PS Claim 27; SEQ ID No 29271; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. AAG47348-AAG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 22 AA;
 XX
 Query Match 1.8%; Score 8; DB 22; Length 22;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDDG 444
 DB 7 DDDDDDDG 14
 XX
 RESULT 48
 AAB83385
 ID AAB83385 standard; peptide; 22 AA.
 XX
 AC AAB83385;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Lunasin fragment #6.
 XX
 KM Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
 KM prostate; liver; kidney; therapy; soybean.
 XX
 OS Glycine max.
 XX
 PN WO200172784-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09453.
 XX
 PR 24-MAR-2000; 2000US-0534705.
 XX
 PA (FILG-) FILGEN BIOSCIENCES INC.
 XX
 PI Galvez AF;
 PT WPI; 2001-648438/74.
 XX
 PT Treatment or prevention of cancer by administering a lunasin peptide
 PT which binds to non-acetylated histones H3 and H4 and prevents their
 PT acetylation in hypacetylated repressed chromatin -
 XX

PS Claim 7; Page 42; 49pp; English.
XX
XX This sequence represents a lunasin peptide used in the method of the
CC invention. The method is for the treatment or prevention of cancer,
CC comprising administration of a lunasin peptide or its active fragment or
CC analog. The invention is used to treat or prevent cancer, particularly
CC in a human. The cancers that can be treated and prevented include
CC those of the colon, upper gastrointestinal tract, breast, prostate,
CC liver, kidney or any other internal organs or tissues.
XX
SQ Sequence 22 AA;
Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDD 443
Db 13 GDDDDDD 20
RESULT 49
ABB30607
ID ABB30607 standard; Peptide; 22 AA.
XX
XX ABB30607;
AC
XX 01-FEB-2002 (first entry)
DT
XX Peptide #3258 encoded by breast cell single exon nucleic acid probe.
XX
XX Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00662.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-496933/54.
DR
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 27; SEQ ID NO 13575; 327pp + sequence listing; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater

CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 22 AA;
Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 7 DDDDDDDG 14
RESULT 50
ABB35771
ID ABB35771 standard; Peptide; 22 AA.
XX
XX ABB35771;
AC
XX 04-FEB-2002 (first entry)
DT
XX
XX Peptide #3277 encoded by human foetal liver single exon probe.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00669.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-483447/52.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PT
XX
XX Claim 27; SEQ ID NO 28406; 639pp + sequence listing; English.
PS
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 22 AA;
Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 7 DDDDDDDG 14

RESULT 51

ABB21195 ID ABB21195 standard; Protein; 22 AA.

AC ABB21195;

DT 23-JAN-2002 (first entry)

DE Protein #3194 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

KM cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

OS Homo sapiens.

XX MO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PS Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

PS Claim 15; SEQ ID NO 22965; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA11305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g., cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 22 AA;

QY Query Match 1.8%; Score 8; DB 22; Length 22;

Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 7 DDDDDDDG 14

RESULT 52

AAM56580 ID AAM56580 standard; Protein; 22 AA.

AC AAM56580;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28685.

XX Human; brain expressed exon; gene expression analysis; probe;

KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer.

OS Homo sapiens.

XX WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PS Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 28685; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX the probes of the invention.

PS Sequence 22 AA;

QY Query Match 1.8%; Score 8; DB 22; Length 22;

Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 7 DDDDDDDG 14

RESULT 53

AAM68960 ID AAM68960 standard; Protein; 22 AA.

AC AAM68960;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29266.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukemia; lymphoma; myeloma.

XX Homo sapiens.

OS

XX MO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488900/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human bone marrow -
XX
XX
PS Example 4; SEQ ID NO: 29266; 658bp + Sequence Listing; English.
XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 7 DDDDDDDG 14

RESULT 54
AAM16787
ID AAM16787 standard; Protein; 22 AA.
XX
XX AAM16787;
AC

DT 12-OCT-2001 (first entry)

DE Peptide #3221 encoded by probe for measuring cervical gene expression.
XX
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157278-A2.
XX
XX PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00670.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488901/53.
XX DR
XX
XX

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX
XX Claim 27; SEQ ID No 21613; 487bp; English.
XX
XX

CC The present invention relates to human single exon nucleic acid probes
CC (SEN: see A110068-A128459). The present sequence is a peptide encoded
CC by one such probe. The SENs are derived from human HeLa cells. The SENs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX

SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 7 DDDDDDDG 14

RESULT 55
AAM29271
ID AAM29271 standard; Protein; 22 AA.
XX
XX AAM29271;
AC

DT 17-OCT-2001 (first entry)

DE Peptide #3308 encoded by probe for measuring placental gene expression.

XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00663.
XX

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-48897/53.
XX DR
XX
XX

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 29540; 654bp; English.
PS

XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI1315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.

XX Sequence 22 AA;
SQ

Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||||
7 DDDDDDDG 14

Db

RESULT 56
AAM04504
ID AAM04504 standard; Protein; 22 AA.
XX
XX AAM04504;
AC
XX
XX 09-OCT-2001 (first entry)
DT
XX
XX Peptide #3186 encoded by probe for measuring breast gene expression.
DE
XX
XX Probe; human; breast disease; breast cancer; development disorder;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
OS
XX
XX MO200157270-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 29-JAN-2001; 2001WO-US000661.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PR
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-476286/51.
DR
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast.
PT
XX
XX Claim 27; SEQ ID No 13244; 322pp; English.
PS
XX
XX The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer; disorders of development,
CC inflammatory diseases of the breast; fibrocystic changes; proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 22 AA;
Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||||
7 DDDDDDDG 14

Db

RESULT 57
AAB62619
ID AAB62619 standard; peptide; 22 AA.
XX
XX AAB62619;
AC
XX
XX 23-JUL-2001 (first entry)
DT
XX
XX Soybean lunasin peptide variant (residues 22-43).
DE
XX
XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.
KM
XX
XX Glycine max.
OS
XX
XX MO200134808-A2.
PN
XX
XX 17-MAY-2001.
PD
XX
XX 12-NOV-2000; 2000WO-US31211.
PF
XX
XX 12-NOV-1999; 99US-0165334.
PR
XX
XX (FILG-) FILGEN BIOSCIENCES INC.
PA
XX
XX Galvez AF;
PI
XX
XX WPI; 2001-343605/36.
DR
XX
XX Large scale production of lunasin, a cancer preventive peptide from
PT soybean, by preparing lunasin gene constructs using Pichia expression
PT vectors, optimizing lunasin expression parameters and purifying lunasin
PT
XX
XX Claim 7; Page 62; 67pp; English.
PS
XX
XX The invention provides a method of producing lunasin, a cancer preventive
CC peptide from soybean, or its active variant by recombinant DNA technology
CC in large quantities. The method comprises preparing lunasin gene
CC constructs or its variants using protein expression vectors, optimizing
CC lunasin expression parameters for large-scale production and performing
CC a series of isolation and purification to obtain large quantities of
CC biologically active recombinant lunasin peptide. The lunasin peptide
CC exhibits inhibitory effects against malignant transformation of cells
CC induced by chemical carcinogens and viral oncogenes. Sequences
CC AAB62614-24 represent soybean lunasin variants.
CC
XX
XX Sequence 22 AA;
SQ

Query Match 1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
|||||
13 GDDDDDD 20

Db

RESULT 58
ABG38546
ID ABG38546 standard; Peptide; 22 AA.
XX
XX ABG38546;

XX 19-AUG-2002 (first entry)
 DT Human peptide encoded by genome-derived single exon probe SEQ ID 28211.
 XX
 DE Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS
 XX
 PN WO200186003-A2.
 XX
 XX 15-NOV-2001.
 PD
 XX
 XX 30-JAN-2001; 2001WO-US00665.
 PF
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 PT WPI; 2002-114183/15.
 PS
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 measure gene expression in human lung samples -
 Claim 27; SEQ ID No 28211; 634bp; English.

CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 22 AA;
 QY
 DB 437 DDDDDDDG 444
 7 DDDDDDDG 14
 Query Match 1.8%; Score 8; DB 23; Length 22;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 59
 AAB83395
 ID AAB83395 standard; peptide; 25 AA.
 XX
 AC AAB83395;
 XX
 DT 26-MAR-2002 (first entry)
 DE
 XX Lunasin related peptide #4.
 XX
 KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
 KW prostate; liver; kidney; therapy; soybean.
 XX
 OS Glycine max.
 XX
 PN WO200172784-A2.
 PD
 XX 04-OCT-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09453.
 PF
 XX 24-MAR-2000; 2000US-0534705.
 PR
 XX (FILG-) FILGEN BIOSCIENCES INC.
 XX
 PI Galvez AF;
 XX
 DR WPI; 2001-648438/74.
 XX
 PT Treatment or prevention of cancer by administering a lunasin peptide
 PT which binds to non-acetylated histones H3 and H4 and prevents their
 PT acetylation in hypoacetylated repressed chromatin -
 XX
 PS Disclosure; Page 47; 49pp; English.
 XX
 XX This sequence represents a lunasin related peptide used in the method of
 CC the invention. The method is for the treatment or prevention of cancer,
 CC comprising administration of a lunasin peptide or its active fragment or
 CC analog. The invention is used to treat or prevent cancer, particularly
 CC in a human. The cancers that can be treated and prevented include
 CC those of the colon, upper gastrointestinal tract, breast, prostate,
 CC liver, kidney or any other internal organs or tissues.
 XX
 SQ Sequence 25 AA;
 QY
 DB 436 GDDDDDDD 443
 16 GDDDDDDD 23
 Query Match 1.8%; Score 8; DB 22; Length 25;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 60
ABG49409
ID ABG49409 standard; Peptide; 31 AA.
XX
AC ABG49409;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID No 28057.
XX
KM Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KM hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
PS Claim 27; SEQ ID No 28057; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 31 AA;
XX
Query Match 1.8%; Score 8; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 16 DDDDDDDG 23

```

```

XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #2059 encoded by breast cell single exon nucleic acid probe.
XX
KM Human; microarray; single exon probe; gene expression; breast;
KM disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 27; SEQ ID No 12376; 327bp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 31 AA;
XX
Query Match 1.8%; Score 8; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 16 DDDDDDDG 23

```

```

RESULT 61
ABB29408
ID ABB29408 standard; Peptide; 31 AA.
XX
AC ABB29408;
XX

```

```

RESULT 62
ABB34590
ID ABB34590 standard; Peptide; 31 AA.
XX
AC ABB34590;
XX

```

DT 04-FEB-2002 (first entry)
 XX Peptide #2096 encoded by human foetal liver single exon probe.
 XX
 DE Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 KM Homo sapiens.
 OS
 XX MO200157277-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 PS Claim 27, SEQ ID NO 27225; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 31 AA;
 QY
 Db 437 DDDDDDDG 444
 16 DDDDDDDG 23
 Query Match 1.8%; Score 8; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-48899/53.
 DR
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PS Claim 15; SEQ ID NO 21766; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 31 AA;
 QY
 Db 437 DDDDDDDG 444
 16 DDDDDDDG 23
 Query Match 1.8%; Score 8; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 64
 ID AAM55379
 AC AAM55379;
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27484.
 XX
 KM Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 OS
 XX Homo sapiens.
 XX
 PN WO200157275-A2.
 PD
 XX 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 27484; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX Sequence 31 AA;
SQ
Query Match 1.8%; Score 8; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 16 DDDDDDDG 23
RESULT 65
AAM67774
ID AAM67774 standard; Protein; 31 AA.
XX
AC AAM67774;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 28080.
DE
XX Human bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX WO200157276-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
PR
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 28080; 658bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX Sequence 31 AA;
SQ
Query Match 1.8%; Score 8; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
Db 16 DDDDDDDG 23
RESULT 66
AAM15583
ID AAM15583 standard; Protein; 31 AA.
XX
XX AAM15583;
XX
XX 12-OCT-2001 (first entry)
DT
XX Peptide #2017 encoded by probe for measuring cervical gene expression.
DE
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer.
XX
XX Homo sapiens.
OS
XX WO200157278-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
PR
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488901/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID No 20409; 487bp; English.
PS
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs; see AAM10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 31 AA;
SQ
Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
16 DDDDDDDG 23

RESULT 67

AMM28075
ID AAM28075 standard; Protein; 31 AA.

AC AAM28075;

DT 17-OCT-2001 (first entry)

DE Peptide #2112 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder.

XX Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 28344; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SNP;

CC see A11315-A157546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for

CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.

XX Sequence 31 AA;

QY Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
16 DDDDDDDG 23

RESULT 68

AMM03327

ID AAM03327 standard; Protein; 31 AA.

AC AAM03327;

DT 09-OCT-2001 (first entry)

XX Peptide #2009 encoded by probe for measuring breast gene expression.

DE Probe; human; breast disease; breast cancer; development disorder;

KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US00661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression

PT in a human breast -

PS Claim 27; SEQ ID No 12067; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes

CC (see A100010-A110067). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for measuring human gene expression in

CC a human breast sample, where the probe hybridises at high stringency to a

CC nucleic acid expressed in the human breast. The probes are useful for

CC predicting, diagnosing, grading, staging, monitoring and prognosing

CC diseases of the human breast, particularly those diseases with polygenic

CC aetiology. The diseases include: breast cancer; disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 31 AA;

QY Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
16 DDDDDDDG 23

RESULT 69

ABG37322

ID ABG37322 standard; Peptide; 31 AA.

AC ABG37322;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 26987.

KW Human; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 XX MO200186003-A2.
 XX 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 XX Claim 27; SEQ ID NO 26987; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 XX from human lung comprising single exon nucleic acid probes having one of
 XX 12614 nucleic acid sequences mentioned in the specification, or their
 XX complements or the 12387 open reading frames derived from the 12614
 XX probes. Also included are a microarray comprising the novel set of
 XX probes; the novel set of probes which hybridize at high stringency to a
 XX nucleic acid expressed in the human lung; measuring gene expression in a
 XX sample derived from human lung, comprising (a) contacting the array with
 XX a collection of detectably labeled nucleic acids derived from human lung
 XX mRNA, and (b) measuring the label detectably bound to each probe of
 XX the array; identifying exons in a eukaryotic genome, comprising
 XX (a) algorithmically predicting at least one exon from genomic sequences
 XX of the eukaryote; and (b) detecting specific hybridization of detectably
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 XX having a fragment identical to the predicted exon, the probe is included
 XX in the above mentioned microarray; assigning exons to a single gene,
 XX comprising (a) identifying exons from genomic sequence by the method
 XX above and (b) measuring the expression of each of the exons in several
 XX tissues and/or cell types using hybridization to a single exon
 XX microarrays having a probe with the exon, where a common pattern of
 XX expression of the exons in the tissues and/or cell types indicates that
 XX the exons should be assigned to a single gene; a peptide comprising one
 XX of 12011 sequences, mentioned in the specification, or encoded by the
 XX probes/open reading frames (ORF). The probes are used for gene
 XX expression analysis, and for identifying exons in a gene, particularly
 XX using human lung derived mRNA and for the study of lung diseases
 XX such as asthma, lung cancer, chronic obstructive pulmonary disease
 XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 XX haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 XX pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 XX and hyaline membrane disease. The present sequence is a peptide/protein
 XX encoded by a single exon probe of the invention.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.

SQL Sequence 31 AA;
 Query Match 1.8%; Score 8; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 437 DDDDDDDG 444
 DB 16 DDDDDDDG 23
 RESULT 70
 AAB83381
 ID AAB83381 standard; peptide; 41 AA.
 AC AAB83381;
 XX 26-MAR-2002 (first entry)
 DE Lunasin fragment #2.
 XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
 KW prostate; liver; kidney; therapy; soybean.
 XX Glycine max.
 OS Glycine max.
 XX MO200172784-A2.
 XX 04-OCT-2001.
 XX 23-MAR-2001; 2001WO-US09453.
 XX 24-MAR-2000; 2000US-0534705.
 XX (FIIIG-) FIIIGEN BIOSCIENCES INC.
 XX Galvez AF;
 XX WPI; 2001-648438/74.
 XX Treatment or prevention of cancer by administering a lunasin peptide
 XX PT which binds to non-acetylated histones H3 and H4 and prevents their
 XX PT acetylation in hypoacetylated repressed chromatin -
 XX Claim 7; Page 41; 49pp; English.
 XX This sequence represents a lunasin peptide used in the method of the
 XX invention. The method is for the treatment or prevention of cancer,
 XX comprising administration of a lunasin peptide or its active fragment or
 XX analog. The invention is used to treat or prevent cancer, particularly
 XX in a human. The cancers that can be treated and prevented include
 XX those of the colon, upper gastrointestinal tract, breast, prostate,
 XX liver, kidney or any other internal organs or tissues.
 XX Sequence 41 AA;
 Query Match 1.8%; Score 8; DB 22; Length 41;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 436 GDDDDDDDD 443
 DB 34 GDDDDDDDD 41
 RESULT 71
 AAB62615
 ID AAB62615 standard; peptide; 41 AA.
 AC AAB62615;
 XX 23-JUL-2001 (first entry)

DE	Soybean lunasin peptide variant (residues 1-41).
XX	
KM	lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.
XX	
OS	Glycine max.
PX	
PN	WO200134808-A2.
XX	
XX	17-MAY-2001.
PD	
XX	
PF	12-NOV-2000; 2000WO-US31211.
XX	
PR	12-NOV-1999; 99US-0165334.
XX	
PA	(FILG-) FILGEN BIOSCIENCES INC.
XX	
P1	Galvez AF;
XX	
DR	WPI; 2001-343605/36.
XX	
PT	LARGE scale production of lunasin, a cancer preventive peptide from soybean, by preparing lunasin gene constructs using pichia expression vectors, optimizing lunasin expression parameters and purifying lunasin
PT	-
PS	Claim 7; Page 60; 67pp; English.
XX	
CC	The invention provides a method of producing lunasin, a cancer preventive peptide from soybean, or its active variant by recombinant DNA technology in large quantities. The method comprises preparing lunasin gene constructs or its variants using protein expression vectors, optimizing lunasin expression parameters for large-scale production and performing a series of isolation and purification to obtain large quantities of biologically active recombinant lunasin peptide. The lunasin peptide exhibits inhibitory effects against malignant transformation of cells induced by chemical carcinogens and viral oncogenes. Sequences AA62614-24 represent soybean lunasin variants.
CC	
XX	Sequence 41 AA:
SQ	
Query Match	1.8%; Score 8; DB 22; Length 41;
Best Local Similarity	100.0%; Pred. No. 10;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	436 GDDDDDDD 443
DB	34 GDDDDDDD 41
RESULT 72	
ABG51319	
ID	ABG51319 standard; Peptide; 42 AA.
XX	
AC	ABG51319;
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human liver peptide, SEQ ID NO 29967.
XX	
KW	Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200157273-A2.
XX	
PD	
XX	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00664.
XX	
RR	04-FEB-2000; 2000US-0180312.
XX	
PR	26-MAY-2000; 2000US-0207456.
XX	
PR	30-JUN-2000; 2000US-0608408.
XX	

PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
PA	(MOLE-) MOLECULAR DYNAMICS INC.
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
DR	WPI; 2001-488898/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human adult liver -
PS	Claim 27; SEQ ID No 29967; 658bp; English.
XX	
CC	The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC	measuring human gene expression in a sample derived from human adult
CC	liver comprising one of 13109 defined nucleotide sequences given in the
CC	specification (or complements/ fragments). The probe hybridises at high
CC	stringency to a nucleic acid molecule expressed in the human adult
CC	liver. (I) may be used for predicting, measuring and displaying gene
CC	expression in samples derived from human adult liver. The genes
CC	identified may be involved in genetic liver diseases such as cirrhosis,
CC	hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC	is associated with coronary heart disease. ABG47348-ABG59930 represent
CC	human liver single exon encoded peptides of the invention.
CC	Note: The sequence information for this patent does not appear in the
CC	printed specification but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 42 AA:
Query Match	1.8%; Score 8; DB 22; Length 42;
Best Local Similarity	100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	437 DDDDDDDG 444
Db	14 DDDDDDDG 21
RESULT 73	
AAAB83380	
ID	AAAB83380 standard; peptide; 42 AA.
XX	
AC	AAAB83380;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Lunasin fragment #1.
XX	
KM	Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
KM	prostate; liver; kidney; therapy; soybean.
XX	
OS	Glycine max.
XX	
PV	WO200172784-A2.
XX	
PD	04-OCT-2001.
XX	
PF	23-MAR-2001; 2001WO-US09453.
XX	
PR	24-MAR-2000; 2000US-0534705.
PA	(FTIG-) FTIGEN BIOSCIENCES INC.
XX	
PI	Galvez AF;
XX	
DR	WPI; 2001-648438/74.
XX	
PT	Treatment or prevention of cancer by administering a lunasin peptide
PT	which binds to non-acetylated histones H3 and H4 and prevents their

PT acetylation in hypoacetylated repressed chromatin -
 XX
 XX
 PS Claim 7; Page 40; 49pp; English.

CC This sequence represents a lunasin peptide used in the method of the
 CC invention. The method is for the treatment or prevention of cancer,
 CC comprising administration of a lunasin peptide or its active fragment or
 CC analog. The invention is used to treat or prevent cancer, particularly
 CC in a human. The cancers that can be treated and prevented include
 CC those of the colon, upper gastrointestinal tract, breast, prostate,
 CC liver, kidney or any other internal organs or tissues.

XX
 SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
 DB 34 GDDDDDD 41

RESULT 74

ABB31264
 ID ABB31264 standard; Peptide; 42 AA.

XX
 AC ABB31264;

DT 01-FEB-2002 (first entry)

XX Peptide #3915 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;
 KM disease; cancer.

XX Homo sapiens.

XX WO200157271-A2.

XX
 PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -

XX Claim 27; SEQ ID NO 14232; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery and for
 CC determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
 DB 14 DDDDDDDG 21

RESULT 75

ABB36472
 ID ABB36472 standard; Peptide; 42 AA.

XX
 AC ABB36472;

DT 04-FEB-2002 (first entry)

XX Peptide #3978 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX
 PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 29107; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
Db 14 DDDDDDDG 21

Search completed: December 5, 2003, 09:30:17
Job time : 49 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:31:43 ; Search time 38 Seconds
(without alignments)
2226.910 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSLSGAFYQIKFDLQFPE.....GDDDDDDGDEEDNDNDNSE 455

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 684280 seqs, 185983659 residues

Word size : 0

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 100 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	455	100.0	455	US-09-757-982-5	Sequence 5, Appl1
2	455	100.0	455	US-10-094-749-2477	Sequence 2477, Ap
3	269	59.1	349	US-10-106-698-6345	Sequence 6345, Ap
4	80	17.6	92	US-09-764-868-799	Sequence 799, App
5	9	2.0	502	US-09-903-068-12	Sequence 12, Appl
6	9	2.0	856	US-10-231-778-221	Sequence 221, Appl
7	9	2.0	911	US-10-115-482-50	Sequence 48, Appl
8	9	2.0	915	US-10-115-482-48	Sequence 48, Appl
9	9	2.0	1818	US-10-335-687A-2	Sequence 2, Appl1
10	9	2.0	1824	US-10-335-687A-5	Sequence 5, Appl1
11	8	1.8	21	US-10-252-256-3	Sequence 3, Appl1
12	8	1.8	22	US-09-864-761-36493	Sequence 36493, A
13	8	1.8	25	US-10-252-256-5	Sequence 5, Appl1
14	8	1.8	31	US-09-864-761-35294	Sequence 35294, A
15	8	1.8	42	US-09-864-761-37113	Sequence 37113, A

16	1.8	43	15	US-10-252-256-1	Sequence 1, Appl1
17	1.8	180	12	US-09-836-392-34	Sequence 34, Appl
18	1.8	149	10	US-10-032-585-7073	Sequence 7073, Ap
19	1.8	207	12	US-10-225-810-28	Sequence 28, Appl
20	1.8	236	10	US-09-809-545A-6	Sequence 6, Appl1
21	1.8	262	12	US-10-032-585-7551	Sequence 7551, Ap
22	1.8	275	12	US-10-032-585-7271	Sequence 7271, Ap
23	1.8	286	10	US-09-801-368-188	Sequence 188, App
24	1.8	302	12	US-10-029-386-33745	Sequence 33745, A
25	1.8	328	10	US-09-862-027-18	Sequence 18, Appl
26	1.8	393	12	US-10-032-585-7155	Sequence 7155, Ap
27	1.8	416	10	US-09-862-027-14	Sequence 14, Appl1
28	1.8	416	11	US-09-291-417-5	Sequence 5, Appl1
29	1.8	416	12	US-10-353-690-116	Sequence 116, App
30	1.8	416	15	US-10-177-223-107	Sequence 307, App
31	1.8	448	12	US-10-225-810-44	Sequence 44, Appl
32	1.8	487	9	US-09-810-808-8	Sequence 8, Appl1
33	1.8	491	12	US-09-822-110-2	Sequence 2, Appl1
34	1.8	518	15	US-10-283-023-2	Sequence 2, Appl1
35	1.8	534	12	US-09-998-491-2	Sequence 2, Appl1
36	1.8	579	12	US-10-384-743-4	Sequence 4, Appl1
37	1.8	579	12	US-10-136-895-4	Sequence 4, Appl1
38	1.8	590	14	US-10-384-743-15	Sequence 15, Appl1
39	1.8	590	14	US-10-158-895-15	Sequence 15, Appl1
40	1.8	732	12	US-10-032-585-7546	Sequence 7546, Ap
41	1.8	746	15	US-10-153-668-232	Sequence 232, App
42	1.8	746	15	US-10-231-778-225	Sequence 225, App
43	1.8	746	15	US-10-231-778-226	Sequence 226, App
44	1.8	847	14	US-10-143-133-2	Sequence 2, Appl1
45	1.8	938	15	US-10-205-823-64	Sequence 64, Appl1
46	1.8	945	12	US-10-032-585-7114	Sequence 7114, Ap
47	1.8	1221	12	US-10-032-585-7602	Sequence 7602, Ap
48	1.8	1324	12	US-10-203-311A-6	Sequence 6, Appl1
49	1.8	1318	12	US-10-200-562-197	Sequence 197, App
50	1.8	1418	12	US-10-237-551-197	Sequence 197, App
51	1.5	8	11	US-09-882-291-22	Sequence 22, Appl1
52	1.5	9	11	US-09-882-291-29	Sequence 29, Appl1
53	1.5	12	11	US-09-876-904A-173	Sequence 173, App
54	1.5	12	11	US-10-177-550-23	Sequence 23, Appl
55	1.5	14	11	US-09-911-569-80	Sequence 80, Appl
56	1.5	14	12	US-10-200-879-80	Sequence 80, Appl
57	1.5	15	12	US-10-350-405-221	Sequence 221, App
58	1.5	20	9	US-09-864-761-38908	Sequence 38908, A
59	1.5	20	12	US-10-192-832-44	Sequence 44, Appl
60	1.5	20	12	US-10-192-832-45	Sequence 45, Appl
61	1.5	20	12	US-10-192-832-48	Sequence 48, Appl
62	1.5	20	12	US-10-192-832-49	Sequence 49, Appl
63	1.5	20	12	US-10-192-832-50	Sequence 50, Appl
64	1.5	20	12	US-10-192-832-51	Sequence 51, Appl
65	1.5	23	9	US-09-864-761-36028	Sequence 36028, A
66	1.5	24	9	US-09-864-761-34620	Sequence 34620, A
67	1.5	24	9	US-09-864-761-36142	Sequence 36142, A
68	1.5	26	9	US-09-864-761-35447	Sequence 35447, A
69	1.5	27	9	US-09-864-761-39392	Sequence 39392, A
70	1.5	30	11	US-09-911-569-87	Sequence 87, Appl
71	1.5	30	12	US-10-200-879-87	Sequence 87, Appl
72	1.5	30	12	US-10-192-832-58	Sequence 58, Appl
73	1.5	30	12	US-10-192-832-59	Sequence 59, Appl
74	1.5	30	12	US-10-192-832-62	Sequence 62, Appl
75	1.5	30	12	US-10-192-832-63	Sequence 63, Appl
76	1.5	30	12	US-10-192-832-64	Sequence 64, Appl
77	1.5	30	12	US-10-192-832-65	Sequence 65, Appl
78	1.5	30	12	US-10-192-832-72	Sequence 72, Appl
79	1.5	30	12	US-10-192-832-74	Sequence 74, Appl
80	1.5	30	12	US-10-192-832-75	Sequence 75, Appl
81	1.5	30	15	US-10-200-659-3	Sequence 3, Appl1
82	1.5	31	9	US-09-864-761-33555	Sequence 33555, A
83	1.5	34	9	US-09-864-761-44046	Sequence 44046, A
84	1.5	36	9	US-09-864-761-35933	Sequence 35933, A
85	1.5	37	9	US-09-864-761-43116	Sequence 43116, A
86	1.5	38	15	US-10-300-616-34	Sequence 34, Appl
87	1.5	40	9	US-09-864-761-39102	Sequence 39102, A
88	1.5	40	15	US-10-252-256-2	Sequence 2, Appl1

```
89 7 1.5 42 9 US-09-864-761-35018 Sequence 35018, A
90 7 1.5 42 9 US-09-864-761-35629 Sequence 35629, A
91 7 1.5 44 12 US-10-029-386-31572 Sequence 31672, A
92 7 1.5 44 12 US-10-012-697-1554 Sequence 1554, A
93 7 1.5 46 15 US-10-062-548-125 Sequence 125, A
94 7 1.5 51 9 US-09-864-761-39869 Sequence 39869, A
95 7 1.5 61 9 US-09-864-761-39079 Sequence 39079, A
96 7 1.5 81 15 US-10-028-247-2 Sequence 2, A
97 7 1.5 86 14 US-10-028-247-2 Sequence 2, A
98 7 1.5 91 10 US-09-873-880-4 Sequence 4, A
99 7 1.5 98 12 US-10-092-947A-61 Sequence 61, A
100 7 1.5 100 12 US-10-092-947A-57 Sequence 57, A
```

ALIGNMENTS

```
RESULT 1
US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US2002094559A1
; GENERAL INFORMATION:
; APPLICANT: Accion, Suean
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/757, 982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163, 115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 100.0%; Score 455; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKMTISQDKEVAVKKLTKEAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKMTISQDKEVAVKKLTKEAEILSVL 60
QY 61 SHRNIIQFYGVILPEPNYGVITEYASLGSLYDYINSRSEEMDMHIMWATDVAKGMHY 120
Db 61 SHRNIIQFYGVILPEPNYGVITEYASLGSLYDYINSRSEEMDMHIMWATDVAKGMHY 120
QY 121 LHMEAPVKVYIHRDLKSRNVVIADGVKLCIDFGASRFHNTTMSLVGTFPMMAPEVIQS 180
Db 121 LHMEAPVKVYIHRDLKSRNVVIADGVKLCIDFGASRFHNTTMSLVGTFPMMAPEVIQS 180
QY 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSFAELH 240
Db 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSFAELH 240
QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNSTLHNKAEMRCIEATLERLKLKLERD 300
Db 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNSTLHNKAEMRCIEATLERLKLKLERD 300
QY 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLLPLAARMSEESYFESKTESNSAEMSCQI 360
Db 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLLPLAARMSEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSKRGRK 420
Db 361 TATSNGEGHGMNPSLOAMMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDPDLSEGGDDDDGDEEDNDMDNSE 455
Db 421 KVNMAIGFSDPDLSEGGDDDDGDEEDNDMDNSE 455
```

```
RESULT 2
US-10-094-749-2477
; Sequence 2477, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, MACHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094, 749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350, 435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2477
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2477

Query Match 100.0%; Score 455; DB 12; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKMTISQDKEVAVKKLTKEAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKMTISQDKEVAVKKLTKEAEILSVL 60
QY 61 SHRNIIQFYGVILPEPNYGVITEYASLGSLYDYINSRSEEMDMHIMWATDVAKGMHY 120
Db 61 SHRNIIQFYGVILPEPNYGVITEYASLGSLYDYINSRSEEMDMHIMWATDVAKGMHY 120
QY 121 LHMEAPVKVYIHRDLKSRNVVIADGVKLCIDFGASRFHNTTMSLVGTFPMMAPEVIQS 180
Db 121 LHMEAPVKVYIHRDLKSRNVVIADGVKLCIDFGASRFHNTTMSLVGTFPMMAPEVIQS 180
QY 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSFAELH 240
Db 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPSFAELH 240
QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNSTLHNKAEMRCIEATLERLKLKLERD 300
Db 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNSTLHNKAEMRCIEATLERLKLKLERD 300
QY 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLLPLAARMSEESYFESKTESNSAEMSCQI 360
Db 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLLPLAARMSEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSKRGRK 420
Db 361 TATSNGEGHGMNPSLOAMMLMGFGDIFSNKAKAGAVHSGMQINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDPDLSEGGDDDDGDEEDNDMDNSE 455
```


Db 421 KVMALGFSDFDLSEGGDDDDDDGDEEDNDMDNSE 455

RESULT 3
US-10-106-698-6345
; Sequence 6345, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO: 6345
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (340)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6345

Query Match 59.1%; Score 269; DB 15; Length 349;
Best Local Similarity 100.0%; Pred. No. 2,4e-243;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFYQIKRDDIOFENCGSGSGSYRAKWTISQDKEVAVKLLIKTEKEATISLV 60
Db 57 MSSLGASFYQIKRDDIOFENCGSGSGSYRAKWTISQDKEVAVKLLIKTEKEATISLV 116
QY 61 SHRNIIQFVGLIEPPNYGIVTEYASISGLSYDYINSRSEMDMDHIMTWATVAKGMHY 120
Db 117 SHRNIIQFVGLIEPPNYGIVTEYASISGLSYDYINSRSEMDMDHIMTWATVAKGMHY 176
QY 121 LHMEAVKVIHRDLKSRNVIAADGVKICDPSASRFHNTTMSLVGTFFMAPEVIQS 180
Db 177 LHMEAVKVIHRDLKSRNVIAADGVKICDPSASRFHNTTMSLVGTFFMAPEVIQS 236
QY 181 LPSVSECDMYSGVVLVMEMLTREVPKGLGLQVAVLVKNERLTTPSSCPRSPAFELH 240
Db 237 LPSVSECDMYSGVVLVMEMLTREVPKGLGLQVAVLVKNERLTTPSSCPRSPAFELH 296
QY 241 QCWEADAKKRPSEKQIISLESMSNDTSL 269
Db 297 QCWEADAKKRPSEKQIISLESMSNDTSL 325

RESULT 4
US-09-764-868-799
; Sequence 799, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 799
; LENGTH: 92
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-799

Query Match 17.6%; Score 80; DB 10; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.5e-67;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 AEILSVLSHRNIIQFVGLIEPPNYGIVTEYASISGLSYDYINSRSEMDMDHIMTWATD 113
Db 13 AEILSVLSHRNIIQFVGLIEPPNYGIVTEYASISGLSYDYINSRSEMDMDHIMTWATD 72
QY 114 VAKGMHYLHMEAPVKVIHRD 133
Db 73 VAKGMHYLHMEAPVKVIHRD 92

RESULT 5
US-09-903-068-12
; Sequence 12, Application US/09903068
; Patent No. US20020123139A1
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohsei, Dijke, Peter Ten;
; Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
; Having Serine Threonine Kinase Domains And Their Use
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pelfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/903, 068
; FILING DATE: 11-Jul-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/679,187
; FILING DATE: <Unknown>
; APPLICATION NUMBER: PCT/GB93/02367
; FILING DATE: 17-No. US20020123139A1ember-1993
; APPLICATION NUMBER: 9224057.1
; FILING DATE: 17-No. US20020123139A1ember-1992
; APPLICATION NUMBER: 9304677.9
; FILING DATE: 8-March-1993
; APPLICATION NUMBER: 9304680.3
; FILING DATE: 8-March-1993
; APPLICATION NUMBER: 9311047.6
; FILING DATE: 28-May-1993
; APPLICATION NUMBER: 9313763.6
; FILING DATE: 2-July-1993
; APPLICATION NUMBER: 9136099.2
; FILING DATE: 3-August-1993
; APPLICATION NUMBER: 9321344.5
; FILING DATE: 15-October-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohsei, Vaneet
; REGISTRATION NUMBER: 37, 003
; REFERENCE/DOCKET NUMBER: LUD 5298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-903-068-12

Query Match 2.0%; Score 9; DB 10; Length 502;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
Db 327 HRDLKSRNV 335

RESULT 6

US-10-231-778-221
Sequence 221, Application US/10231778
Publication No. US20030126647A1
GENERAL INFORMATION:
APPLICANT: Biodeau, Pierre
APPLICANT: Chaudhury, Abdul M.
APPLICANT: Dennis, Elizabeth S.
APPLICANT: Koltunow, Anna M.G.
APPLICANT: Luo, Ming
APPLICANT: Peacock, William J.
TITLE OF INVENTION: Method for inducing seed development by down-regulating
TITLE OF INVENTION: expression of the F1S2 gene
FILE REFERENCE: 72-98A
CURRENT APPLICATION NUMBER: US/10/231,778
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: 09/398,237
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/101,184
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: AU PP6061
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6062
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6063
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU P01345
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: AU P01346
PRIOR FILING DATE: 1999-07-01
NUMBER OF SEQ ID NOS: 239
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 221
LENGTH: 856
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: EZAI peptide
OTHER INFORMATION: fragment
US-10-231-778-221

Query Match 2.0%; Score 9; DB 15; Length 856;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
Db 17 DDDDDGEE 25

RESULT 7

US-10-115-482-50
Sequence 50, Application US/10115482
Publication No. US20030212257A1
GENERAL INFORMATION:
APPLICANT: Spytek, et al.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
TITLE OF INVENTION: AND METHODS

TITLE OF INVENTION: OF USING THE SAME
FILE REFERENCE: 21404-322D
CURRENT APPLICATION NUMBER: US/10/115,482
CURRENT FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/285,325
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: 60/285,890
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/286,068
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/286,292
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/287,213
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/288,257
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/291,134
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/291,725
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/294,771
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/296,965
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: 60/299,128
PRIOR FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 149
SEQ ID NO 50
LENGTH: 911
TYPE: PRT
ORGANISM: Homo sapiens
US-10-115-482-50

Query Match 2.0%; Score 9; DB 12; Length 911;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSYRA 33
Db 282 GSFGSYRA 290

RESULT 8
US-10-115-482-48
Sequence 48, Application US/10115482
Publication No. US20030212257A1
GENERAL INFORMATION:
APPLICANT: Spytek, et al.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
TITLE OF INVENTION: AND METHODS
FILE REFERENCE: 21404-322D
CURRENT APPLICATION NUMBER: US/10/115,482
CURRENT FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863

PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/285,325
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: 60/285,890
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/286,068
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/286,292
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/287,213
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/288,257
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/291,134
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/291,725
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/294,771
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/296,965
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: 60/299,128
PRIOR FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 149
SEQ ID NO 48
LENGTH: 915
TYPE: PRT
ORGANISM: Homo sapiens
US-10-115-482-48

Query Match 2.0%; Score 9; DB 12; Length 915;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GSGSVYRA 33
DB 282 GSGSVYRA 290
RESULT 9
US-10-335-687A-2
Sequence 2, Application US/10335687A
Publication No. US20030166222A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: 39267, Human Kinase Family Members and
FILE REFERENCE: MP102-001PIRNM
CURRENT APPLICATION NUMBER: US/10/335,687A
CURRENT FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: 60/345,773
PRIOR FILING DATE: 2002-01-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1818
TYPE: PRT
ORGANISM: Homo sapiens
US-10-335-687A-2

Query Match 2.0%; Score 9; DB 12; Length 1818;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSVYRA 33
DB 1185 GSGSVYRA 1193

RESULT 10
US-10-335-687A-5
Sequence 5, Application US/10335687A
Publication No. US20030166222A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: 39267, Human Kinase Family Members and
FILE REFERENCE: MP102-001PIRNM
CURRENT APPLICATION NUMBER: US/10/335,687A
CURRENT FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: 60/345,773
PRIOR FILING DATE: 2002-01-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1824
TYPE: PRT
ORGANISM: Homo sapiens
US-10-335-687A-5

Query Match 2.0%; Score 9; DB 12; Length 1824;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSVYRA 33
DB 1185 GSGSVYRA 1193

RESULT 11
US-10-252-256-3
Sequence 3, Application US/10252256
Publication No. US20030027765A1
GENERAL INFORMATION:
APPLICANT: GALVEZ, ALFREDO F.
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
CANCER THERAPY
FILE REFERENCE: 3729.02
CURRENT APPLICATION NUMBER: US/10/252,256
CURRENT FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: 60/165,334
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 21
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: trlunasin
US-10-252-256-3

Query Match 1.8%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
DB 12 GDDDDDD 19

RESULT 12
US-09-864-761-36493
Sequence 36493, Application US/09864761
Patent No. US20020048763A1

```

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36493
LENGTH: 22
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011235.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1
US-09-864-761-36493

Query Match      1.8%; Score 8; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      437 DDDDDDDG 444
      |||||
      7 DDDDDDDG 14

```

```

RESULT 13
US-10-252-256-5
Sequence 5, Application US/10252256
Publication No. US20030027765A1
GENERAL INFORMATION:
APPLICANT: GALVEZ, ALFREDO F.
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
CANCER THERAPY
FILE REFERENCE: 3729.02
CURRENT APPLICATION NUMBER: US/10/252,256
CURRENT FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: 60/165,334
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 25
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: NLS-trlunasin
US-10-252-256-5

Query Match      1.8%; Score 8; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      436 GDDDDDD 443
      |||||
      16 GDDDDDD 23

RESULT 14
US-09-864-761-35294
Sequence 35294, Application US/09864761
Patent No. US20020648763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 35294
LENGTH: 31
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011235.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
US-09-864-761-35294

Query Match 1.8%; Score 8; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
Db 16 DDDDDDDG 23

RESULT 15
US-09-864-761-37113
Sequence 37113 Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmlca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37113
LENGTH: 42
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO A1229043.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
OTHER INFORMATION: EST_HUMAN HIT: AW863068.1, EVALUE 3.00e-07
US-09-864-761-37113

Query Match 1.8%; Score 8; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
Db 14 DDDDDDDG 21

RESULT 16
US-10-252-256-1
Sequence 1 Application US/10252256
Publication No. US20030027765A1
GENERAL INFORMATION:
APPLICANT: GALVEZ, ALFREDO F.
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
CANCER THERAPY
FILE REFERENCE: 3729.02
CURRENT APPLICATION NUMBER: US/10/252,256
PRIOR FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: 60/165,334
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 43
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Lunasin
US-10-252-256-1

Query Match 1.8%; Score 8; DB 15; Length 43;

Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443
Db 34 GDDDDDDD 41

RESULT 17
US-09-836-392-34

; Sequence 34, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, a

; FILE REFERENCE: PTO20P1

; CURRENT APPLICATION NUMBER: US/09/836,392

; PRIORITY FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: PCT/US00/28066

; PRIOR FILING DATE: 2000-10-11

; PRIOR APPLICATION NUMBER: 60/159,542

; PRIOR FILING DATE: 1999-10-15

; PRIOR APPLICATION NUMBER: 60/165,914

; PRIOR FILING DATE: 1999-11-17

; PRIOR APPLICATION NUMBER: 60/189,027

; PRIOR FILING DATE: 2000-03-14

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 49

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-836-392-34

Query Match 1.8%; Score 8; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 9.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
Db 30 VLKICDFG 37

RESULT 18
US-10-032-585-7073

; Sequence 7073, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7073

; LENGTH: 180

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7073

Query Match 1.8%; Score 8; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 30;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443
Db 117 GDDDDDDD 124

RESULT 19
US-10-225-810-28

; Sequence 28, Application US/10225810

; Publication No. US20030157512A1

; GENERAL INFORMATION:

; APPLICANT: Birmingham, Jr., John R.

; TITLE OF INVENTION: Transdorsin and Methods of Using Transdorsin

; FILE REFERENCE: McLaugh-07165

; CURRENT APPLICATION NUMBER: US/10/225,810

; CURRENT FILING DATE: 2002-08-21

; NUMBER OF SEQ ID NOS: 76

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 28

; LENGTH: 207

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-225-810-28

Query Match 1.8%; Score 8; DB 12; Length 207;
Best Local Similarity 100.0%; Pred. No. 34;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 15 WMAPEVIQ 22

RESULT 20
US-09-809-545A-6

; Sequence 6, Application US/09809545A

; Patent No. US20020110804A1

; GENERAL INFORMATION:

; APPLICANT: Stanton, Lawrence W.

; APPLICANT: White, R. Tyler

; TITLE OF INVENTION: SECRETED FACTORS

; FILE REFERENCE: SCIOS.017A

; CURRENT APPLICATION NUMBER: US/09/809,545A

; CURRENT FILING DATE: 2001-03-14

; NUMBER OF SEQ ID NOS: 84

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 236

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-809-545A-6

Query Match 1.8%; Score 8; DB 10; Length 236;
Best Local Similarity 100.0%; Pred. No. 38;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443
Db 41 GDDDDDDD 48

RESULT 21
US-10-032-585-7551

; Sequence 7551, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7551

; LENGTH: 262

TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7551

Query Match 1.8%; Score 8; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 43 DDDDDDDG 50

RESULT 22
US-10-032-585-7271
Sequence 7271, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7271
LENGTH: 275
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7271

Query Match 1.8%; Score 8; DB 12; Length 275;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443
DB 259 GDDDDDDD 266

RESULT 23
US-09-801-368-188
Sequence 188, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 188
LENGTH: 286

TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-188

Query Match 1.8%; Score 8; DB 10; Length 286;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443
DB 108 GDDDDDDD 115

RESULT 24
US-10-029-386-33745
Sequence 33745, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33745
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ACO23066.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 15
OTHER INFORMATION: SWISSPROT HIT: Q01105, EVALU1 1.00e-111
US-10-029-386-33745

Query Match 1.8%; Score 8; DB 12; Length 302;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDD 444
DB 257 DDDDDDDD 264

RESULT 25
US-09-862-027-18
Sequence 18, Application US/09862027
Patent No. US20020142428A1
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/345,473
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 328
TYPE: PRT
ORGANISM: C. elegans
US-09-862-027-18

Query Match 1.8%; Score 8; DB 10; Length 328;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEATLIS 58
|||||

Db 90 EKEAEILS 97

RESULT 26
US-10-032-585-7155
; Sequence 7155, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Giang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Busey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7155
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7155

Query Match 1.8%; Score 8; DB 12; Length 393;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 266 GDDDDDD 273

RESULT 27
US-09-862-027-14
; Sequence 14, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-14

Query Match 1.8%; Score 8; DB 10; Length 416;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 28
US-09-291-417-5
; Sequence 5, Application US/09291417A
; Publication No. US20030050230A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 240/300
; CURRENT APPLICATION NUMBER: US/09/291,417A

CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: US 60/081,784
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mammalian (Human) STLK2
US-09-291-417-5

Query Match 1.8%; Score 8; DB 11; Length 416;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 29
US-10-353-690-116
; Sequence 116, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Myoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodrigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 33394, 3484, 345, 9252, 935,
; TITLE OF INVENTION: 10532, 16610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MP102-018PROMINIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-10-353-690-116

Query Match 1.8%; Score 8; DB 12; Length 416;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 30

US-10-177-293-307
; Sequence 307, Application US/10177293
; Publication No. US20030124126A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatc, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamakkar, Shubhangt
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzsrai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 307
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-307

Query Match 1.8%; Score 8; DB 15; Length 416;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 31

US-10-225-810-44
; Sequence 44, Application US/10225810
; Publication No. US20030157512A1
; GENERAL INFORMATION:

; APPLICANT: Bermingham, Jr., John R.
; TITLE OF INVENTION: Transdomin and Methods of Using Transdomin
; FILE REFERENCE: McLaugh-07165
; CURRENT APPLICATION NUMBER: US/10/225,810
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-225-810-44

Query Match 1.8%; Score 8; DB 12; Length 448;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 189 WMAPEVIQ 196

RESULT 32
US-09-810-808-8
; Sequence 8, Application US/09810808
; Patent No. US20020042114A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; Guegler, Karl J.
; Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/810,808
; FILING DATE: 15-Mar-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/541,228
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1117791
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-810-808-8

Query Match 1.8%; Score 8; DB 9; Length 487;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
 DB 190 WMAPEVIQ 197

RESULT 33

US-09-822-110-2
 ; Sequence 2, Application US/09822110
 ; Publication No. US20030170252A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Hwa-Chain R.
 ; TITLE OF INVENTION: ANTIBODY COMPOSITIONS SPECIFIC FOR p33QIK AND p63KRS1
 ; FILE REFERENCE: 4350.000800
 ; CURRENT APPLICATION NUMBER: US/09/822,110
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 60/193,550
 ; PRIOR FILING DATE: 2000-03-31
 ; NUMBER OF SEQ ID NOS: 76
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 491
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-822-110-2

Query Match 1.8%; Score 8; DB 12; Length 491;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
 DB 187 WMAPEVIQ 194

RESULT 34

US-10-283-023-2
 ; Sequence 2, Application US/10283023
 ; Publication No. US20030091573A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Carroll, Joseph M.
 ; TITLE OF INVENTION: Methods and compositions for the
 ; TITLE OF INVENTION: diagnosis and treatment of hematological disorders using
 ; FILE REFERENCE: MP101-239PIRM
 ; CURRENT APPLICATION NUMBER: US/10/283,023
 ; PRIOR FILING DATE: 2002-10-29
 ; PRIOR APPLICATION NUMBER: 60/335,044
 ; PRIOR FILING DATE: 2001-10-31
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 518
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-10-283-023-2

Query Match 1.8%; Score 8; DB 15; Length 518;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
 DB 170 VLKICDFG 177

RESULT 35

US-09-998-491-2
 ; Sequence 2, Application US/09998491
 ; Publication No. US20030165528A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mileusnic, Radmila

APPLICANT: Rose, Stephen Peter Russell
 ; TITLE OF INVENTION: Polypeptides and their uses
 ; FILE REFERENCE: 3578-120
 ; CURRENT APPLICATION NUMBER: US/09/998,491
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: GB 0109558.7
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: GB 0120084
 ; PRIOR FILING DATE: 2001-08-07
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 534
 ; TYPE: PRT
 ; ORGANISM: Chick
 ; US-09-998-491-2

Query Match 1.8%; Score 8; DB 12; Length 534;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDG 444
 DB 84 DDDDDDG 91

RESULT 36

US-10-384-743-4
 ; Sequence 4, Application US/10384743
 ; Publication No. US20030162228A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ONO, KOICHIRO
 ; APPLICANT: OHTOMO, TOSHIHIKO
 ; APPLICANT: TSUCHIYA, MASAYUKI
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
 ; FILE REFERENCE: 053466/0278
 ; CURRENT APPLICATION NUMBER: US/10/384,743
 ; CURRENT FILING DATE: 2003-03-11
 ; PRIOR APPLICATION NUMBER: US/09/529,279
 ; PRIOR FILING DATE: 2000-04-11
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796
 ; PRIOR FILING DATE: 1998-10-22
 ; PRIOR APPLICATION NUMBER: JP 9/290188
 ; PRIOR FILING DATE: 1997-10-22
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-384-743-4

Query Match 1.8%; Score 8; DB 12; Length 579;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
 DB 170 VLKICDFG 177

RESULT 37

US-10-158-895-4
 ; Sequence 4, Application US/10158895
 ; Publication No. US20020155624A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ONO, KOICHIRO
 ; APPLICANT: OHTOMO, TOSHIHIKO
 ; APPLICANT: TSUCHIYA, MASAYUKI
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
 ; FILE REFERENCE: 053466/0278
 ; CURRENT APPLICATION NUMBER: US/10/158,895
 ; CURRENT FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: US/09/529,279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 579
TYPE: PRT
ORGANISM: Homo sapiens
US-10-158-895-4

Query Match
Best Local Similarity 100.0%; Score 8; DB 14; Length 579;
Pred. No. 84;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
Db 170 VLKICDFG 177

RESULT 38
US-10-384-743-15
Sequence 15, Application US/10384743
Publication No. US20030162228A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/384,743
CURRENT FILING DATE: 2003-03-11
PRIOR APPLICATION NUMBER: US/09/529,279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-384-743-15

Query Match
Best Local Similarity 100.0%; Score 8; DB 12; Length 590;
Pred. No. 85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
Db 170 VLKICDFG 177

RESULT 39
US-10-158-895-15
Sequence 15, Application US/10158895
Publication No. US20020155624A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158,895
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529,279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796

PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-158-895-15

Query Match
Best Local Similarity 100.0%; Score 8; DB 14; Length 590;
Pred. No. 85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153
Db 170 VLKICDFG 177

RESULT 40
US-10-032-585-7546
Sequence 7546, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jitang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7546
LENGTH: 732
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7546

Query Match
Best Local Similarity 100.0%; Score 8; DB 12; Length 732;
Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 106 GDDDDDD 113

RESULT 41
US-10-153-668-232
Sequence 232, Application US/10153668
Publication No. US20030092616A1
GENERAL INFORMATION:
APPLICANT: HONDA, Goichi
APPLICANT: MATSUDA, Akio
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAWA, Kenya
TITLE OF INVENTION: STAT6 Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/293,172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316,031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30

;; PRIOR APPLICATION NUMBER: JP 2001-313175
;; PRIOR FILING DATE: 2001-10-10
;; NUMBER OF SEQ ID NOS: 488
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 232
;; LENGTH: 746
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-153-668-232

Query Match 1.8%; Score 8; DB 15; Length 746;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
Db 183 DDDDDDDG 190

RESULT 42
US-10-231-778-225
;; Sequence 225, Application US/10231778
;; Publication No. US20030126647A1
;; GENERAL INFORMATION:
;; APPLICANT: Billoreau, Pierre
;; APPLICANT: Dennis, Elizabeth S.
;; APPLICANT: Koltunow, Anna M.G.
;; APPLICANT: Luo, Ming
;; APPLICANT: Peacock, William J.
;; TITLE OF INVENTION: Method for inducing seed development by down-regulating
;; FILE REFERENCE: 72-98A
;; CURRENT APPLICATION NUMBER: US/10/231,778
;; PRIOR FILING DATE: 2002-11-08
;; PRIOR APPLICATION NUMBER: 09/398,237
;; PRIOR FILING DATE: 1999-09-20
;; PRIOR APPLICATION NUMBER: 60/101,184
;; PRIOR FILING DATE: 1998-09-21
;; PRIOR APPLICATION NUMBER: AU PP6061
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU PP6062
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU PP6063
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU P01345
;; PRIOR FILING DATE: 1999-07-01
;; PRIOR APPLICATION NUMBER: AU P01346
;; PRIOR FILING DATE: 1999-07-01
;; NUMBER OF SEQ ID NOS: 239
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 225
;; LENGTH: 746
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: EZH2 peptide
US-10-231-778-225

Query Match 1.8%; Score 8; DB 15; Length 746;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
Db 183 DDDDDDDG 190

RESULT 43
US-10-231-778-226
;; Sequence 226, Application US/10231778
;; Publication No. US20030126647A1

;; GENERAL INFORMATION:
;; APPLICANT: Billoreau, Pierre
;; APPLICANT: Chaudhury, Abdul M.
;; APPLICANT: Dennis, Elizabeth S.
;; APPLICANT: Koltunow, Anna M.G.
;; APPLICANT: Luo, Ming
;; APPLICANT: Peacock, William J.
;; TITLE OF INVENTION: Method for inducing seed development by down-regulating
;; FILE REFERENCE: 72-98A
;; CURRENT APPLICATION NUMBER: US/10/231,778
;; PRIOR FILING DATE: 2002-11-08
;; PRIOR APPLICATION NUMBER: 09/398,237
;; PRIOR FILING DATE: 1999-09-20
;; PRIOR APPLICATION NUMBER: 60/101,184
;; PRIOR FILING DATE: 1998-09-21
;; PRIOR APPLICATION NUMBER: AU PP6061
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU PP6062
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU PP6063
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: AU P01345
;; PRIOR FILING DATE: 1999-07-01
;; PRIOR APPLICATION NUMBER: AU P01346
;; PRIOR FILING DATE: 1999-07-01
;; NUMBER OF SEQ ID NOS: 239
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 226
;; LENGTH: 746
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Ezh1 peptide
US-10-231-778-226

Query Match 1.8%; Score 8; DB 15; Length 746;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
|||
Db 183 DDDDDDDG 190

RESULT 44
US-10-143-133-2
;; Sequence 2, Application US/10143133
;; Publication No. US20020197658A1
;; GENERAL INFORMATION:
;; APPLICANT: Yoganathan, Thillainathan
;; APPLICANT: Delaney, Allen
;; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
;; FILE REFERENCE: KINE-023
;; CURRENT APPLICATION NUMBER: US/10/143,133
;; PRIOR FILING DATE: 2002-05-09
;; PRIOR APPLICATION NUMBER: 60/290,555
;; PRIOR FILING DATE: 2001-05-10
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 847
;; TYPE: PRT
;; ORGANISM: Homo sapien
US-10-143-133-2

Query Match 1.8%; Score 8; DB 14; Length 847;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRDLKS 136
|||||

DB 237 VIHRLKS 244

RESULT 45

US-10-205-823-64

; Sequence 64, Application US/10205823

; Publication No. US20030108963A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Endege, Wilson O.

; APPLICANT: Gannavarapu, Manjula

; APPLICANT: Gorbacheva, Belia

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Womsey, Angela M.

; APPLICANT: Zhao, Xumei

; APPLICANT: Anderson, Dustin

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MRI-044

; CURRENT APPLICATION NUMBER: US/10/205,823

; CURRENT FILING DATE: 2002-07-25

; PRIOR APPLICATION NUMBER: 60/307,982

; PRIOR FILING DATE: 2001-07-25

; PRIOR APPLICATION NUMBER: 60/314,356

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/325,020

; PRIOR FILING DATE: 2001-09-25

; PRIOR APPLICATION NUMBER: 60/341,746

; PRIOR FILING DATE: 2001-12-12

; PRIOR APPLICATION NUMBER: 60/362,158

; NUMBER OF SEQ ID NOS: 455

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 64

; LENGTH: 938

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-205-823-64

Query Match 1.8%; Score 8; DB 15; Length 938;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDDD 441

DB 598 SEGDDDDD 605

RESULT 46

US-10-032-585-7114

; Sequence 7114, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7114

; LENGTH: 945

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7114

Query Match 1.8%; Score 8; DB 12; Length 945;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443

DB 914 GDDDDDDD 921

RESULT 47

US-10-032-585-7602

; Sequence 7602, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7602

; LENGTH: 1221

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7602

Query Match 1.8%; Score 8; DB 12; Length 1221;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDGGE 445

DB 873 DDDDDGGE 880

RESULT 48

US-10-203-311A-6

; Sequence 6, Application US/10203311A

; Publication No. US20030186321A1

; GENERAL INFORMATION:

; APPLICANT: PHARMA PACIFIC

; APPLICANT: Meitert, Jean Francois

; APPLICANT: Dron, Michel

; APPLICANT: Tovey, Michael Gerard

; TITLE OF INVENTION: INTERPERON-ALPHA INDUCED GENE

; FILE REFERENCE: 46658/250044

; CURRENT APPLICATION NUMBER: US/10/203,311A

; CURRENT FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: GB 0002979.3

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: GB 0002980.1

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: GB 0002982.7

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: GB 0002981.9

; PRIOR FILING DATE: 2000-02-09

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 1234

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-203-311A-6

Query Match 1.8%; Score 8; DB 12; Length 1234;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443

Db 1222 GDDDDDD 1229

RESULT 49
US-10-200-562-197
; Sequence 197, Application US/10200562
; Publication No. US20030165819A1
; GENERAL INFORMATION:
; APPLICANT: McGowan, Patrick
; APPLICANT: Hosken, Nancy A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C2
; CURRENT APPLICATION NUMBER: US/10/200,562
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: HSV2
US-10-200-562-197

Query Match 1.8%; Score 8; DB 12; Length 1318;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 444
Db 614 DDDDDDD 621

RESULT 50
US-10-237-551-197
; Sequence 197, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: HSV2
US-10-237-551-197

Query Match 1.8%; Score 8; DB 12; Length 1318;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 444
Db 614 DDDDDDD 621

RESULT 51
US-09-882-291-22
; Sequence 22, Application US/09882291
; Publication No. US20030040472A1
; GENERAL INFORMATION:
; APPLICANT: Zealand Pharmaceuticals A/S
; TITLE OF INVENTION: NO. US20030040472A1 Peptide Conjugates
; FILE REFERENCE: 007-2001
; CURRENT APPLICATION NUMBER: US/09/882,291
; CURRENT FILING DATE: 2001-06-15

NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence
US-09-882-291-22

Query Match 1.5%; Score 7; DB 11; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443
Db 2 DDDDDDD 8

RESULT 52
US-09-882-291-29
; Sequence 29, Application US/09882291
; Publication No. US20030040472A1
; GENERAL INFORMATION:
; APPLICANT: Zealand Pharmaceuticals A/S
; TITLE OF INVENTION: NO. US20030040472A1 Peptide Conjugates
; FILE REFERENCE: 007-2001
; CURRENT APPLICATION NUMBER: US/09/882,291
; CURRENT FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence
US-09-882-291-29

Query Match 1.5%; Score 7; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 53
US-09-876-904A-173
; Sequence 173, Application US/09876904A
; Publication No. US2003002794A1
; GENERAL INFORMATION:
; APPLICANT: BOULIKAS, TENI
; TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIPOGENES TM) AND THERAPEUTIC
; TITLE OF INVENTION: AGENTS WITH NUCLEAR LOCALIZATION SIGNAL/FUSOGENIC PEPTIDE
; FILE REFERENCE: TB-2002.00
; CURRENT APPLICATION NUMBER: US/09/876,904A
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,925
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 173
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Karyophilic peptide
US-09-876-904A-173

Query Match 1.5%; Score 7; DB 11; Length 12;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLK 135
|||||

Db 5 VHRDLK 11

RESULT 54

US-10-177-550-23
; Sequence 23, Application US/10177550
; Publication No. US2003013933A1
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: No. US2003013933A1e1 Integrin Alpha Subunit
; FILE REFERENCE: L00560/70009(MAT/TAU)
; CURRENT APPLICATION NUMBER: US/10/177,550
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 09/293,238
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 08/879,338
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: US 08/663,731
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: US 08/199,776
; PRIOR FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-177-550-23

Query Match 1.5%; Score 7; DB 12; Length 12;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
|||||

Db 3 DDDDDD 9

RESULT 55

US-09-911-569-80
; Sequence 80, Application US/09911569
; Publication No. US20030069173A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NEILSON, PAMELA
; APPLICANT: LAN, JIANQING
; APPLICANT: SHIH, POJEN
; APPLICANT: JESSE, JOEL A.
; APPLICANT: SCHIFFERLI, KEVIN P.
; APPLICANT: GEBREYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/911,569

FILING DATE: 23-Jul-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 14

OTHER INFORMATION: /product="OTHER"

/note="G AT POSITION 14 CAN BE ABSENT"

SEQUENCE DESCRIPTION: SEQ ID NO: 80:

Query Match 1.5%; Score 7; DB 11; Length 14;

Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDG 444
|||||

Db 8 DDDDDG 14

RESULT 56

US-10-200-879-80
; Sequence 80, Application US/10200879
; Publication No. US20030144230A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY-NEILSON, PAMELA
; APPLICANT: LAN, JIANQING
; APPLICANT: SHIH, POJEN
; APPLICANT: JESSE, JOEL A.
; APPLICANT: SCHIFFERLI, KEVIN P.
; APPLICANT: GEBREYEHU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/200,879
; FILING DATE: 23-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/911,569
; FILING DATE: 23-Jul-2001
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; APPLICATION NUMBER: US 08/818,200

FILING DATE: 14-MAR-1997
APPLICATION NUMBER: US 08/658,130
FILING DATE: 04-JUN-1996
APPLICATION NUMBER: US 08/477,354
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: SULLIVAN, SALLY A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 32-95E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: Modified-site
LOCATION: 14
OTHER INFORMATION: /product= "OTHER"
/note= "G AT POSITION 14 CAN BE ABSENT"
US-10-200-879-80
SEQUENCE DESCRIPTION: SEQ ID NO: 80:

Query Match 1.5%; Score 7; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDG 444
Db 8 DDDDDG 14

RESULT 57
US-10-350-405-221
Sequence 221, Application US/10350405
Publication No. US20030215894A1
GENERAL INFORMATION:
APPLICANT: Niman, Henry L.
TITLE OF INVENTION: Polypeptide-Induced Monoclonal Receptors
FILE REFERENCE: TSRI 35.5 CON 7/LIG
CURRENT APPLICATION NUMBER: US/10350,405
PRIOR FILING DATE: 2003-01-24
PRIOR APPLICATION NUMBER: 09/427,576
PRIOR FILING DATE: 1999-10-26
PRIOR APPLICATION NUMBER: 08/461,583
PRIOR FILING DATE: 1995-06-02
PRIOR APPLICATION NUMBER: 08/294,879
PRIOR FILING DATE: 1994-08-23
PRIOR APPLICATION NUMBER: 08/054,864
PRIOR FILING DATE: 1993-04-28
PRIOR APPLICATION NUMBER: 07/900,502
PRIOR FILING DATE: 1992-06-16
PRIOR APPLICATION NUMBER: 07/780,415
PRIOR FILING DATE: 1991-10-22
NUMBER OF SEQ ID NOS: 227
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 221
LENGTH: 15
TYPE: PRT
ORGANISM: retrovirus
US-10-350-405-221

Query Match 1.5%; Score 7; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178
Db 5 WMAPEVI 11

RESULT 58
US-09-864-761-38908
Sequence 38908, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1a-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/532,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38908
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004633.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
US-09-864-761-38908

Query Match 1.5%; Score 7; DB 9; Length 20;

Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 13 DDDDDD 19

RESULT 59

US-10-192-832-44
; Sequence 44, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUGUANG
; APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTV-043.01
; CURRENT APPLICATION NUMBER: US/10/192,832
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/304,256
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-44

Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 1 DDDDDD 7

RESULT 60

US-10-192-832-45
; Sequence 45, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUGUANG
; APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTV-043.01
; CURRENT APPLICATION NUMBER: US/10/192,832
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/304,256
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-45

Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 11 DDDDDD 17

RESULT 61

US-10-192-832-48
; Sequence 48, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUGUANG
; APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTV-043.01
; CURRENT APPLICATION NUMBER: US/10/192,832
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/304,256
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-48

Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 1 DDDDDD 7

RESULT 62

US-10-192-832-49
; Sequence 49, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUGUANG
; APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTV-043.01
; CURRENT APPLICATION NUMBER: US/10/192,832
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/304,256
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-49

Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 11 DDDDDD 17

RESULT 63

US-10-192-832-50
; Sequence 50, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUGUANG

```
APPLICANT: VAUTHEY, SYLVAIN
TITLE OF INVENTION: SUPRACANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTW-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 50
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-50
```

```
Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
```

```
QY 437 DDDDDD 443
DB 1 DDDDDD 7
```

```
RESULT 64
US-10-192-832-51
Sequence 51, Application US/10192832
Publication No. US20030176335A1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAUTHEY, SYLVAIN
TITLE OF INVENTION: SUPRACANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTW-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 51
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-51
```

```
Query Match 1.5%; Score 7; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
```

```
QY 437 DDDDDD 443
DB 11 DDDDDD 17
```

```
RESULT 65
US-09-864-761-36028
Sequence 36028, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
```

```
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36028
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011309.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3e+02
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-09-864-761-36028
```

```
Query Match 1.5%; Score 7; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
```

```
QY 437 DDDDDD 443
DB 3 DDDDDD 9
```

```
RESULT 66
US-09-864-761-34620
Sequence 34620, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34620
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009491.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; US-09-864-761-34620

Query Match      1.5%; Score 7; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      437 DDDDDD 443
Db      5 DDDDDD 11

RESULT 67
US-09-864-761-36142
```

```

; Sequence 36142, Application US/09864761
; Patent No. US0020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36142
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007159.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.8
; US-09-864-761-36142

Query Match      1.5%; Score 7; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      437 DDDDDD 443
```

Db 1 DDDDDDD 7

```
RESULT 68
US-09-864-761-35447
; Sequence 35447, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35447
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031661.16 SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.3
US-09-864-761-35447
```

Query Match 1.5%; Score 7; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

```
RESULT 69
US-09-864-761-39392
; Sequence 39392, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39392
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005959.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.8
```

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EST_HUMAN HIT: T95763.1, EVALU2 2.10e+00
US-09-864-761-33932

Query Match 1.5%; Score 7; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 70
US-09-911-569-87
Sequence 87, Application US/09911569
Publication No. US20030069173A1
GENERAL INFORMATION:
APPLICANT: HAMLEY-NELSON, PAMELA
LAN, JIANQING
SHIH, POJEN
JESSE, JOEL A.
SCHIFFERLI, KEVIN P.
GEBEYERU, GULILAT
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER & SULLIVAN
STREET: 5370 MANHATTAN CIRCLE, SUITE 201
CITY: BOULDER
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/911,569
FILING DATE: 23-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/039,780
FILING DATE: 16-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: SULLIVAN, SALLY A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 32-95D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Modified-site
LOCATION: 30
OTHER INFORMATION: /product= "OTHER"
/note= "G AT POSITION 30 CAN BE ABSENT"
SEQUENCE DESCRIPTION: SEQ ID NO: 87:
US-09-911-569-87

Query Match 1.5%; Score 7; DB 11; Length 30;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDG 444
Db 24 DDDDDG 30

RESULT 71
US-10-200-879-87
Sequence 87, Application US/10200879
Publication No. US20030144230A1
GENERAL INFORMATION:
APPLICANT: HAMLEY-NELSON, PAMELA
LAN, JIANQING
SHIH, POJEN
JESSE, JOEL A.
SCHIFFERLI, KEVIN P.
GEBEYERU, GULILAT
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER & SULLIVAN
STREET: 5370 MANHATTAN CIRCLE, SUITE 201
CITY: BOULDER
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/200,879
FILING DATE: 23-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/911,569
FILING DATE: 23-JUL-2001
APPLICATION NUMBER: US 09/039,780
FILING DATE: 16-MAR-1998
APPLICATION NUMBER: US 08/818,200
FILING DATE: 14-MAR-1997
APPLICATION NUMBER: US 08/658,130
FILING DATE: 04-JUN-1996
APPLICATION NUMBER: US 08/477,354
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: SULLIVAN, SALLY A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 32-95E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Modified-site
LOCATION: 30
OTHER INFORMATION: /product= "OTHER"
/note= "G AT POSITION 30 CAN BE ABSENT"
SEQUENCE DESCRIPTION: SEQ ID NO: 87:
US-10-200-879-87

Query Match 1.5%; Score 7; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	438	DDDDDDG	444
Db	24	DDDDDDG	30

```

RESULT 72
US-10-192-832-58
? Sequence: 58, Application US/10192832
? Publication No: US2003017633SA1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAOTHEY, SYLVAIN
? TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
? FILE REFERENCE: MTW-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIORITY APPLICATION NUMBER: 60/304,256
PRIORITY FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 58
? LENGTH: 30
? TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-58

```

```

RESULT 73
US-10-192-832-59
Sequence: 59, Application US/10192832
Publication No. US20030176335A1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAITHEY, SYLVAIN
TITLE OF INVENTION: SURFPEPTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTU-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 59
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-59

```

RESULT 74

US-10-192-832-62
Sequence 62. Application US/10192832
Publication No. US2003017635A1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAUTHEY, SYLVAIN
TITLE OF INVENTION: SUPRACACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTW-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 62
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-62

```

RESULT 75
US-10-192-832-63
; Sequence 63. Application US/10192832
; Publication No. US20030176335A1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTU-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 63
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-192-832-63

```

Search completed: December 5, 2003, 09:37:18
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:29:22 ; Search time 21 Seconds
(Without alignments)
916.735 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSLSGAFVQIKFDLQFFE.....GDDDDDDGGEHNDMDNSE 455

Scoring table: OLIGO

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCFUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455	100.0	455	3	US-09-221-235-5
2	455	100.0	455	3	US-09-221-928-5
3	455	100.0	455	3	US-09-221-527-5
4	455	100.0	455	3	US-09-221-236-5
5	455	100.0	455	3	US-09-221-416-5
6	455	100.0	455	3	US-09-221-245-5
7	455	100.0	455	3	US-09-163-115-5
8	455	100.0	455	3	US-09-221-528-5
9	455	100.0	455	3	US-09-593-553-5
10	455	100.0	455	3	US-09-221-237-5
11	455	100.0	455	4	US-09-399-588-2
12	14	3.1	14	1	US-08-375-300-4
13	10	2.2	764	1	US-08-177-431-4
14	10	2.2	764	5	PCT-US95-16930-4
15	10	2.2	1089	1	US-08-375-300-2
16	10	2.2	1089	1	US-09-177-431-2
17	10	2.2	1089	5	PCT-US95-16930-2
18	10	2.2	303	3	US-08-158-735A-15
19	9	2.0	502	3	US-09-382-256-12
20	9	2.0	502	3	US-09-385-115-12
21	9	2.0	502	4	US-08-436-265-12
22	9	2.0	502	4	US-09-679-187-12
23	9	2.0	505	1	US-08-149-105-16
24	9	2.0	505	1	US-08-317-847-16
25	9	2.0	856	4	US-09-699-266A-13
26	9	2.0	2391	2	US-08-446-855A-2
27	9	2.0	2391	2	US-08-446-855A-2

28	9	2.0	2391	3	US-09-150-741-2	Sequence 2, Appl
29	8	1.8	36	1	US-08-240-712-35	Sequence 35, Appl
30	8	1.8	36	1	US-08-443-890-35	Sequence 35, Appl
31	8	1.8	43	4	US-09-303-814-1	Sequence 1, Appl
32	8	1.8	158	2	US-08-618-911-4	Sequence 4, Appl
33	8	1.8	158	2	US-08-618-911-6	Sequence 6, Appl
34	8	1.8	158	4	US-08-938-675A-2	Sequence 2, Appl
35	8	1.8	158	4	US-09-531-727-2	Sequence 2, Appl
36	8	1.8	179	2	US-08-152-132-2	Sequence 2, Appl
37	8	1.8	192	3	US-09-382-080-1	Sequence 1, Appl
38	8	1.8	192	3	US-08-859-937-1	Sequence 1, Appl
39	8	1.8	226	1	US-08-431-080-26	Sequence 26, Appl
40	8	1.8	226	1	US-08-938-534-26	Sequence 26, Appl
41	8	1.8	226	4	US-09-345-294-26	Sequence 26, Appl
42	8	1.8	251	2	US-08-469-537A-80	Sequence 80, Appl
43	8	1.8	251	2	US-08-469-537A-81	Sequence 81, Appl
44	8	1.8	265	4	US-09-252-991A-28073	Sequence 28073, A
45	8	1.8	270	2	US-08-852-743-5	Sequence 5, Appl
46	8	1.8	270	3	US-09-185-370-5	Sequence 5, Appl
47	8	1.8	317	4	US-09-485-077A-17	Sequence 17, Appl
48	8	1.8	328	4	US-09-345-473B-18	Sequence 18, Appl
49	8	1.8	416	2	US-09-211-930-3	Sequence 3, Appl
50	8	1.8	416	2	US-09-211-930-11	Sequence 11, Appl
51	8	1.8	416	3	US-09-340-993-3	Sequence 3, Appl
52	8	1.8	416	3	US-09-340-993-11	Sequence 11, Appl
53	8	1.8	416	4	US-09-468-442-3	Sequence 3, Appl
54	8	1.8	416	4	US-09-468-442-11	Sequence 11, Appl
55	8	1.8	416	4	US-09-685-462-4	Sequence 4, Appl
56	8	1.8	416	4	US-09-345-473B-14	Sequence 14, Appl
57	8	1.8	487	2	US-08-712-709-8	Sequence 8, Appl
58	8	1.8	487	3	US-09-111-444-8	Sequence 8, Appl
59	8	1.8	487	3	US-09-541-228-8	Sequence 8, Appl
60	8	1.8	487	4	US-09-685-462-8	Sequence 8, Appl
61	8	1.8	544	4	US-09-252-991A-26096	Sequence 26096, A
62	8	1.8	579	4	US-09-529-279-4	Sequence 4, Appl
63	8	1.8	579	4	US-10-158-895-4	Sequence 4, Appl
64	8	1.8	590	4	US-09-529-279-15	Sequence 15, Appl
65	8	1.8	590	4	US-10-158-895-15	Sequence 15, Appl
66	8	1.8	943	2	US-08-469-537A-107	Sequence 107, Appl
67	8	1.8	1085	1	US-08-431-080-28	Sequence 28, Appl
68	8	1.8	1085	2	US-08-938-534-28	Sequence 28, Appl
69	8	1.8	1085	4	US-09-345-294-28	Sequence 28, Appl
70	7	1.5	11	2	US-08-482-228-215	Sequence 215, App
71	7	1.5	11	3	US-08-482-528-215	Sequence 215, App
72	7	1.5	12	3	US-08-199-776-23	Sequence 23, Appl
73	7	1.5	12	3	US-08-663-731-23	Sequence 23, Appl
74	7	1.5	12	3	US-08-879-338-23	Sequence 23, Appl
75	7	1.5	12	4	US-09-293-238B-23	Sequence 23, Appl
76	7	1.5	13	4	PCT-US95-0204A-23	Sequence 23, Appl
77	7	1.5	13	4	US-09-625-570-4	Sequence 4, Appl
78	7	1.5	14	4	US-09-039-780A-87	Sequence 87, Appl
79	7	1.5	30	4	US-09-039-780A-87	Sequence 87, Appl
80	7	1.5	30	4	US-09-376-113-3	Sequence 3, Appl
81	7	1.5	38	4	US-09-390-134B-34	Sequence 34, Appl
82	7	1.5	46	4	US-09-369-247-125	Sequence 125, App
83	7	1.5	50	3	US-08-290-736C-9	Sequence 9, Appl
84	7	1.5	73	3	US-08-290-736C-8	Sequence 8, Appl
85	7	1.5	76	2	US-08-117-952-757	Sequence 757, App
86	7	1.5	81	4	US-09-376-113-2	Sequence 2, Appl
87	7	1.5	86	1	US-08-370-225-20	Sequence 20, Appl
88	7	1.5	86	1	US-08-461-859-20	Sequence 20, Appl
89	7	1.5	86	5	PCT-US93-1006B-20	Sequence 20, Appl
90	7	1.5	115	4	US-09-345-236B-11	Sequence 131, App
91	7	1.5	118	1	US-08-072-574-4	Sequence 4, Appl
92	7	1.5	118	1	US-08-486-270-4	Sequence 4, Appl
93	7	1.5	118	3	US-08-367-264-4	Sequence 4, Appl
94	7	1.5	118	4	US-09-153-757-4	Sequence 4, Appl
95	7	1.5	118	4	US-09-459-715-4	Sequence 32007, A
96	7	1.5	119	4	US-09-252-991A-32007	Sequence 5204, Ap
97	7	1.5	123	4	US-09-328-352-5204	Sequence 31, Appl
98	7	1.5	126	4	US-08-483-533-31	Sequence 31, Appl
99	7	1.5	126	4	US-09-283-471A-31	Sequence 12, Appl
100	7	1.5	159	1	US-08-463-090B-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-221-235-5

Sequence 5, Application US/09221235
Patent No. 6043040
GENERAL INFORMATION:
APPLICANT: Accon, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,235
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-235-5

Query Match 100.0%; Score 455; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60
DB 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60
QY 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
DB 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHNEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPREVIO 180
DB 121 LHNEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPREVIO 180
QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEQLQVAMLVKNERLTISSCPSRFAELIH 240
DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEQLQVAMLVKNERLTISSCPSRFAELIH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKKLERD 300
QY 301 LSPFEQELKERERLTKMEQKLTQSNTPLLPLAARMSEESYFESKTESNSAEMSCOI 360
DB 301 LSPFEQELKERERLTKMEQKLTQSNTPLLPLAARMSEESYFESKTESNSAEMSCOI 360
QY 361 TATSNGEGHGMPSLOAMLMGFGDI FSNMKGAVVHSGMOINMOAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMPSLOAMLMGFGDI FSNMKGAVVHSGMOINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 2

US-09-221-928-5
Sequence 5, Application US/09221928
Patent No. 6121030
GENERAL INFORMATION:
APPLICANT: Accon, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115

EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-928-5

Query Match 100.0%; Score 455; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60
DB 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60
QY 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
DB 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHNEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPREVIO 180
DB 121 LHNEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPREVIO 180
QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEQLQVAMLVKNERLTISSCPSRFAELIH 240
DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEQLQVAMLVKNERLTISSCPSRFAELIH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKKLERD 300
QY 301 LSPFEQELKERERLTKMEQKLTQSNTPLLPLAARMSEESYFESKTESNSAEMSCOI 360
DB 301 LSPFEQELKERERLTKMEQKLTQSNTPLLPLAARMSEESYFESKTESNSAEMSCOI 360
QY 361 TATSNGEGHGMPSLOAMLMGFGDI FSNMKGAVVHSGMOINMOAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMPSLOAMLMGFGDI FSNMKGAVVHSGMOINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 3

US-09-221-527-5
Sequence 5, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Accon, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-527-5

Query Match 100.0%; Score 455; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60
DB 1 MSSLGASFWQIKFDDLOFFENCNGGSGSVYRAKWIISQKEVAVKKLTIKKEAEILSVL 60

Qy	61	SHNNIIOFGVILEPPNYGLMEYASLSGLXYIYINSNSEEMDDHIMTATDYAKMNY	120
Db	61	SHNNIIOFGVILEPPNYGLMEYASLSGLXYIYINSNSEEMDDHIMTATDYAKMNY	120
Qy	121	LHNEAPYKVIHRDLKSRNVVIAADGV.LKICDFGASFRHNHTTHMSLVGTFPMAPEVIO	180
Db	121	LHNEAPYKVIHRDLKSRNVVIAADGV.LKICDFGASFRHNHTTHMSLVGTFPMAPEVIO	180
Qy	181	LPIVSECTDYSYGVULMEMULTRBPVKCIBGLQVAMLYVEKNERLITPSCPRSPFAULH	240
Db	181	LPIVSECTDYSYGVULMEMULTRBPVKCIBGLQVAMLYVEKNERLITPSCPRSPFAULH	240
Qy	241	QCEWADAKKRPSEKQIISLESMSNDTSLPDKCSFLFNKAEMRCETIATLERLKKLERD	300
Db	241	QCEWADAKKRPSEKQIISLESMSNDTSLPDKCSFLFNKAEMRCETIATLERLKKLERD	300
Qy	301	LSFKEOELKERERRLMEOKLTEOSNTPLLLPLAARMSSESPSKTEESNSAEMSCQI	360
Db	301	LSFKEOELKERERRLMEOKLTEOSNTPLLLPLAARMSSESPSKTEESNSAEMSCQI	360
Qy	361	TATNSNGBGHMNPSLOAMLMGFQDIFSNKKAAGVMSGMOINMOAKONSKTTSKRGK	420
Db	361	TATNSNGBGHMNPSLOAMLMGFQDIFSNKKAAGVMSGMOINMOAKONSKTTSKRGK	420
Qy	421	KYVMALGFSDFDLSEGGDDDDDDGEEEDNDMNSE	455
Db	421	KYVMALGFSDFDLSEGGDDDDDDGEEEDNDMNSE	455

RESULT 4
US-09-221-236-5
; Sequence 5, Application US/09221236
Defect No. 014084

```

: GENERAL INFORMATION:
: APPLICANT: Acton, Susan
: TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: NMI-050
: CURRENT APPLICATION NUMBER: US/09/221,236
: CURRENT FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: EARLIER FILING DATE: 1998-09-29
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 455
: TYPE: prt
: ORGANISM: Homo sapiens
: US-09-221-236-5

```

Query Match	100.0%;	Score 455;	DB 3;	Length 455;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 455; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	MSLSIASFQOIFEDDLOFENCGGSGFSGVYAKMISQREKVAVKLLIKIEKAEILISYL	60
Db	1	MSLSIASFQIQIFEDDLOFENCGGSGFSGVYAKMISQREKVAVKLLIKIEKAEILISYL	60
QY	61	SHRNIIQFYGVILPEPPNIGIVTEYASLSGLYDINSNRSEMDMHIMTAVDAGMAY	120
Db	61	SHRNIIQFYGVILPEPPNIGIVTEYASLSGLYDINSNRSEMDMHIMTAVDAGMAY	120
QY	121	LHMEAPVKYIHMDLSRNVIADGVLKICDGASGFHNHTTHMSLVGTFPMMAPEVIOS	180
Db	121	LHMEAPVKYIHMDLSRNVIADGVLKICDGASGFHNHTTHMSLVGTFPMMAPEVIOS	180
QY	181	LPVSECTDITYSGVYLMWLTREVPFKGLEGVAVLVEKNERLTI PSSCPSPFAELIH	240
Db	181	LPVSECTDITYSGVYLMWLTREVPFKGLEGVAVLVEKNERLTI PSSCPSPFAELIH	240
QY	241	QWEMDADAKRBPFKQIISILBESMSNDTSLPDKCNSTLHNKAEMRCIEATLELTKLTERD	300
Db	241	QWEMDADAKRBPFKQIISILBESMSNDTSLPDKCNSTLHNKAEMRCIEATLELTKLTERD	300

Qy	301	LSFKBEELKERERRIKOMBEOKLTBESNPILLPLAAMSEBSYESTTEENSASBNSCOI	360
Db	301	LSFKBEELKERERRIKOMBEOKLTBESNPILLPLAAMSEBSYESTTEENSASBNSCOI	360
Qy	361	TATSNQEGHGMPSLQAMMLMGFGDI FSNMAYAGA VMHSGMOINNOAKONSSKTTSKRGK	420
Db	361	TATSNQEGHGMPSLQAMMLMGFGDI FSNMAYAGA VMHSGMOINNOAKONSSKTTSKRGK	420
Qy	421	KVMALGFSDPDLSEGGDDDDDDGEEENDNDINSE	455
Db	421	KVMALGFSDPDLSEGGDDDDDDGEEENDNDINSE	455

RESULT 5

```

Sequence 5. Application US/09221416
Patent No. 6153417
GENERAL INFORMATION:
APPLICANT: Accon, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,416
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-416-5

```

```
Query Match      100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	MSLSIGASFOIIFEDDLOFFENCGGGSFSSVYRAMIISODKEVAAYKLLKIKEKEAILISVL	60
Dp	1	MSLSIGASFOIIFEDDLOFFENCGGGSFSSVYRAMIISODKEVAAYKLLKIKEKEAILISVL	60
Qy	61	SHRNIIIOFYGVILLEPPNIGIVTEYASLSGLSYDIYNSRSEEMDMHIMTAVYAKMTH	120
Dp	61	SHRNIIIOFYGVILLEPPNIGIVTEYASLSGLSYDIYNSRSEEMDMHIMTAVYAKMTH	120
Qy	121	LHNEAPVXKIIHRDLSKRNVVIAADCVLKICOPGASRPHNHTHSLVGTPEPMAPEYIOS	180
Dp	121	LHNEAPVXKIIHRDLSKRNVVIAADCVLKICOPGASRPHNHTHSLVGTPEPMAPEYIOS	180
Qy	181	LPVSEICDYSGYVWLWEMLTREYVFKLEGIQVAMLVENKNERLTPSSCPSRFAELH	240
Dp	181	LPVSEICDYSGYVWLWEMLTREYVFKLEGIQVAMLVENKNERLTPSSCPSRFAELH	240
Qy	241	QCEADAPKRPSFKOIIISLSMSNDISLDPKCNSTFLHNKAEWCEIATLERLAKLERD	300
Dp	241	QCEADAPKRPSFKOIIISLSMSNDISLDPKCNSTFLHNKAEWCEIATLERLAKLERD	300
Qy	301	LSFKEDOLKERERRLKMEOKLTBOSNTPLLPLAARHSEBSYESTTESNSAEMSCOI	360
Dp	301	LSFKEDOLKERERRLKMEOKLTBOSNTPLLPLAARHSEBSYESTTESNSAEMSCOI	360
Qy	361	TATSNCEGHQNNPSIQAMMLMGFGDI FSNMRYAGVMSHGMQINNOAKONSKTTSKRKG	420
Dp	361	TATSNCEGHQNNPSIQAMMLMGFGDI FSNMRYAGVMSHGMQINNOAKONSKTTSKRKG	420
Qy	421	KVNMAIGFSDPDLSEGDODDDDDGEEEDNDNDNE	455
Dp	421	KVNMAIGFSDPDLSEGDODDDDDGEEEDNDNDNE	455

RESULT 6
US-09-221-245-5

```

; Sequence 5, Application US/09221245
; Patent No. 6183958
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER APPLICATION NUMBER: 1998-12-28
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-245-5

```

```

Query Match      100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSLGASFWQIKFDDLOFFENC GGSGFSGVYRAKWI SQDEKVA VVKLLKIEKEAEITLSVL 60
DB 1 MSSLGASFWQIKFDDLOFFENC GGSGFSGVYRAKWI SQDEKVA VVKLLKIEKEAEITLSVL 60
QY 61 SHRNIIQFYGVILPEPNYGI VTEYASL GSYLYDINSNRSEEMDMHIMTATDVAKGMHY 120
DB 61 SHRNIIQFYGVILPEPNYGI VTEYASL GSYLYDINSNRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGV LKICDPGASRFHNHTTHMSLVGTFPMMAPREVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGV LKICDPGASRFHNHTTHMSLVGTFPMMAPREVIOS 180
QY 181 LPVSETCDTYSYGVV LEMLTREVPFKGLEQLQVAMLVVEKNERLTI PSSCPSFAELIHL 240
DB 181 LPVSETCDTYSYGVV LEMLTREVPFKGLEQLQVAMLVVEKNERLTI PSSCPSFAELIHL 240
QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCN SFLHNKAERCEIATELRLKLEERD 300
DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCN SFLHNKAERCEIATELRLKLEERD 300
QY 301 LSPFEOELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
DB 301 LSPFEOELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEGHGNPISLOAMLMFGFDI FSNMKA GAVVHSGMGINNOAKONS KTTSKRGK 420
DB 361 TATSNGEGHGNPISLOAMLMFGFDI FSNMKA GAVVHSGMGINNOAKONS KTTSKRGK 420
QY 421 KVNMA LCFSPDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMA LCFSPDLSEGGDDDDDDGEEEDNDMDNSE 455

```

```

RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-163-115-5

```

```

Query Match      100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSLGASFWQIKFDDLOFFENC GGSGFSGVYRAKWI SQDEKVA VVKLLKIEKEAEITLSVL 60
DB 1 MSSLGASFWQIKFDDLOFFENC GGSGFSGVYRAKWI SQDEKVA VVKLLKIEKEAEITLSVL 60
QY 61 SHRNIIQFYGVILPEPNYGI VTEYASL GSYLYDINSNRSEEMDMHIMTATDVAKGMHY 120
DB 61 SHRNIIQFYGVILPEPNYGI VTEYASL GSYLYDINSNRSEEMDMHIMTATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGV LKICDPGASRFHNHTTHMSLVGTFPMMAPREVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGV LKICDPGASRFHNHTTHMSLVGTFPMMAPREVIOS 180
QY 181 LPVSETCDTYSYGVV LEMLTREVPFKGLEQLQVAMLVVEKNERLTI PSSCPSFAELIHL 240
DB 181 LPVSETCDTYSYGVV LEMLTREVPFKGLEQLQVAMLVVEKNERLTI PSSCPSFAELIHL 240
QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCN SFLHNKAERCEIATELRLKLEERD 300
DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCN SFLHNKAERCEIATELRLKLEERD 300
QY 301 LSPFEOELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
DB 301 LSPFEOELKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEGHGNPISLOAMLMFGDI FSNMKA GAVVHSGMGINNOAKONS KTTSKRGK 420
DB 361 TATSNGEGHGNPISLOAMLMFGDI FSNMKA GAVVHSGMGINNOAKONS KTTSKRGK 420
QY 421 KVNMA LCFSPDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMA LCFSPDLSEGGDDDDDDGEEEDNDMDNSE 455

```

```

RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-528-5

```

QY 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
DB 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
QY 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
DB 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 9
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAKP-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match 100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKMSODKEVAVKKLTIKEAEILLSVL 60
DB 1 MSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKMSODKEVAVKKLTIKEAEILLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLXYDINSNRSEMDMDHMTWATDVAKGMHY 120
DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLXYDINSNRSEMDMDHMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLTCDPGASRFHNHTTHMSLVGTFPMMAPEVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLTCDPGASRFHNHTTHMSLVGTFPMMAPEVIOS 180
QY 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
DB 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
QY 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
DB 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420

QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 10
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAKP-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; EARLIER FILING DATE: 1998-12-28
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-221-237-5

Query Match 100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKMSODKEVAVKKLTIKEAEILLSVL 60
DB 1 MSLGASFVQIKFDDLOFPENCGGSGFSGVYRAKMSODKEVAVKKLTIKEAEILLSVL 60
QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLXYDINSNRSEMDMDHMTWATDVAKGMHY 120
DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLSGLXYDINSNRSEMDMDHMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLTCDPGASRFHNHTTHMSLVGTFPMMAPEVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLTCDPGASRFHNHTTHMSLVGTFPMMAPEVIOS 180
QY 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
DB 181 LPVSECTDYTSYGVMEMLTREVPEKGLGLOVAMLVVEKNERLTIIPSSCPSPFAELLH 240
QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKKLERD 300
QY 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
DB 301 LSFKEBELKERERRLKMWOKLTEQSNTPLLPLAARMSESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSNMKAQAVHSGMQINMOAKONSKTTSKRGRK 420
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 11
US-09-399-588-2
; Sequence 2, Application US/09399588
; Patent No. 6511825
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Rosamaria
; APPLICANT: Callow, Marinella
; TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids
; FILE REFERENCE: 1044-US
; CURRENT APPLICATION NUMBER: US/09/399,588

CURRENT FILING DATE: 1999-09-20
 EARLIER APPLICATION NUMBER: 60/104,088
 EARLIER FILING DATE: 1998-10-13
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 455
 TYPE: PRT
 ORGANISM: Human J42
 US-09-399-588-2

Query Match 100.0%; Score 455; DB 4; Length 455;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKRPDDLOFENCGGSGFSGYRAKWSIQDEKVAVKLLKIEKAEILSVL 60
 DB 1 MSSLGASFWQIKRPDDLOFENCGGSGFSGYRAKWSIQDEKVAVKLLKIEKAEILSVL 60
 QY SHRNIIQFYGVIIIEPPNYGIVTEYASLSGLYDIYNSRSEMDHMTATVAAGMHY 120
 DB SHRNIIQFYGVIIIEPPNYGIVTEYASLSGLYDIYNSRSEMDHMTATVAAGMHY 120
 QY 61 SHRNIIQFYGVIIIEPPNYGIVTEYASLSGLYDIYNSRSEMDHMTATVAAGMHY 120
 DB 61 SHRNIIQFYGVIIIEPPNYGIVTEYASLSGLYDIYNSRSEMDHMTATVAAGMHY 120
 QY 121 LHMEAPKVIHRDLSKNVVIADGVYKICDFGASRPHNTHTMSLVGTFPMNAPEVIQS 180
 DB 121 LHMEAPKVIHRDLSKNVVIADGVYKICDFGASRPHNTHTMSLVGTFPMNAPEVIQS 180
 QY 181 LPVSETCDTYSYGVVLEMLTREVFPKLEGLQVAMLVKNERLTTPSSCPSFAELH 240
 DB 181 LPVSETCDTYSYGVVLEMLTREVFPKLEGLQVAMLVKNERLTTPSSCPSFAELH 240
 QY 241 QCWEADAKKPPSFQIISIESMSNDTSLPDKCNSFLHNAEVRCEIATLERIKLERD 300
 DB 241 QCWEADAKKPPSFQIISIESMSNDTSLPDKCNSFLHNAEVRCEIATLERIKLERD 300
 QY 301 LSKREOKLKERERLKKWEOQLTEQSNTPLLPLAARSESPFESTESNSAEMSCQI 360
 DB 301 LSKREOKLKERERLKKWEOQLTEQSNTPLLPLAARSESPFESTESNSAEMSCQI 360
 QY 361 TATNGEGHGMNPSLQAMMLMGFDIPSMKAGAVHSGQINNOAKONSKTTSKRGK 420
 DB 361 TATNGEGHGMNPSLQAMMLMGFDIPSMKAGAVHSGQINNOAKONSKTTSKRGK 420
 QY 421 KVMNALGFSPPDJSEGGDDDDDEEDNDMDNSE 455
 DB 421 KVMNALGFSPPDJSEGGDDDDDEEDNDMDNSE 455

RESULT 12
 US-09-399-588-5
 Sequence 5, Application US/09399588
 Patent No. 6511825
 GENERAL INFORMATION:
 APPLICANT: Ruggieri, Rosamaria
 APPLICANT: Callow, Marinella
 APPLICANT: Diaz, Paul W.
 TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids
 FILE REFERENCE: 1044-US
 CURRENT APPLICATION NUMBER: US/09/399,588
 EARLIER FILING DATE: 1999-09-20
 EARLIER APPLICATION NUMBER: 60/104,088
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 14
 TYPE: PRT
 ORGANISM: SRK-KA
 US-09-399-588-5

Query Match 3.1%; Score 14; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 5.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 AKONSKTTSKRG 419
 DB 1 AKONSKTTSKRG 14

RESULT 13
 US-08-375-300-4
 Sequence 4, Application US/08375300
 Patent No. 5679566
 GENERAL INFORMATION:
 APPLICANT: Feng, He
 APPLICANT: Jacobson, Allan S.
 TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
 TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECA FUNCTION
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street Suite 3100
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/375,300
 FILING DATE: 20-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Faese, J. P.
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 04020/046001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)542-5070
 TELEFAX: (617)542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 764 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-375-300-4

Query Match 2.2%; Score 10; DB 1; Length 764;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
 DB 577 DDDDDDDGEE 586

RESULT 14
 US-09-177-431-4
 Sequence 4, Application US/09177431
 Patent No. 6071700
 GENERAL INFORMATION:
 APPLICANT: He, Feng
 APPLICANT: Jacobson, Allan S.
 TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
 TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECA FUNCTION
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA

ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-177-431-4

Query Match 2.2%; Score 10; DB 3; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
DB 577 DDDDDDDGEE 586

RESULT 15
PCT-US95-16930-4
Sequence 4, Application PC/TUS9516930
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
PRODUCTION IN THE ABSENCE OF
TITLE OF INVENTION: NONSENSE-MEDIATED MRNA DECAY
TITLE OF INVENTION: FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16930
FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16930-4

Query Match 2.2%; Score 10; DB 5; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
DB 577 DDDDDDDGEE 586

RESULT 16
US-08-375-300-2
Sequence 2, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
THE ABSENCE OF NONSENSE-MEDIATED MRNA DECAY FUNCTION
TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED MRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-375-300-2

Query Match 2.2%; Score 10; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
DB 902 DDDDDDDGEE 911

RESULT 17
US-09-177-431-2
Sequence 2, Application US/09177431
Patent No. 6071700

GENERAL INFORMATION:
APPLICANT: He, Feng
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRICER APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Faase, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-177-431-2

Query Match 2.2%; Score 10; DB 3; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
DB 902 DDDDDDDGEE 911

RESULT 18
PCT-US95-16930-2
Sequence 2, Application PC/TUS9516930
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY
TITLE OF INVENTION: FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16930

FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Faase, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16930-2

Query Match 2.2%; Score 10; DB 5; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446
DB 902 DDDDDDDGEE 911

RESULT 19
US-08-158-735A-15
Sequence 15, Application US/08158735A
Patent No. 6248554
GENERAL INFORMATION:
APPLICANT: COOK, JONATHAN S.
APPLICANT: CORREA, PAUL E.
APPLICANT: KOENIG, BETH B.
APPLICANT: ROSENBAUM, JAN S.
APPLICANT: TING, JERRY
TITLE OF INVENTION: DNA SEQUENCE CODING FOR A BMP RECEPTOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE PROCTOR & GAMBLE COMPANY
STREET: 11810 EAST MIAMI RIVER ROAD
CITY: ROSS
STATE: OH
COUNTRY: USA
ZIP: 45061
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,735A
FILING DATE: 24-NOV-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CORSTANTE, BRAHM J.
REGISTRATION NUMBER: 34,804
REFERENCE/DOCKET NUMBER: 5088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 627-2858
TELEFAX: (513) 627-0260
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-158-735A-15

Query Match 2.0%; Score 9; DB 3; Length 303;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139
Db 128 HRDLKSRNV 136

RESULT 20

US-09-382-256-12

; Sequence 12, Application US/09382256A

; Patent No. 6207814

; GENERAL INFORMATION:

; APPLICANT: MIYAZONO, Kohei

; TEN DIJKE, Peter

; FRANZEN, Petra

; YAMASHITA, Hidetoshi

; HELDIN, Carl-Henrik

; TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS

; HAVING SERINE THREONINE KINASE DOMAINS,

; AND THEIR USE

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 666 Fifth Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/382,256A

; FILING DATE: 24-Aug-1999

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB93/02367

; FILING DATE: No. 6207814ember 17, 1993

; APPLICATION NUMBER: GB 9224057.1

; FILING DATE: No. 6207814ember 17, 1992

; APPLICATION NUMBER: GB 9304677.9

; FILING DATE: March 8, 1993

; APPLICATION NUMBER: GB 9304680.3

; FILING DATE: March 8, 1993

; APPLICATION NUMBER: 9311047.6

; FILING DATE: May 28, 1993

; APPLICATION NUMBER: 9313763.6

; FILING DATE: July 2, 1993

; APPLICATION NUMBER: 9316099.2

; FILING DATE: August 3, 1993

; APPLICATION NUMBER: 321344.5

; FILING DATE: October 15, 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6207814man D. Hanson

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5298.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 318-3000

; TELEFAX: (212) 752-9958

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 502 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-382-256-12

Query Match 2.0%; Score 9; DB 3; Length 502;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139
Db 327 HRDLKSRNV 335

RESULT 21

US-09-395-115-12

; Sequence 12, Application US/09395115

; Patent No. 6271365

; GENERAL INFORMATION:

; APPLICANT: Miyazono, Kohei; DiJke, Peter Ten;

; FRANZEN, Petra; Yamashita, Hidetoshi; Helden, Carl-Henrik

; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felte & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/395,115

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/436,265

; FILING DATE: 30-October-1995

; APPLICATION NUMBER: PCT/GB93/02367

; FILING DATE: 17-No. 6271365ember-1993

; APPLICATION NUMBER: 9224057.1

; FILING DATE: 17-No. 6271365ember-1992

; APPLICATION NUMBER: 9304677.9

; FILING DATE: 8-March-1993

; APPLICATION NUMBER: 9304680.3

; FILING DATE: 8-March-1993

; APPLICATION NUMBER: 9311047.6

; FILING DATE: 28-May-1993

; APPLICATION NUMBER: 9313763.6

; FILING DATE: 2-July-1993

; APPLICATION NUMBER: 9136099.2

; FILING DATE: 3-August-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 9321344.5

; FILING DATE: 15-October-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohler, Vineset

; REGISTRATION NUMBER: 37,003

; REFERENCE/DOCKET NUMBER: LUD 5298

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 502 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-382-256-12

US-09-395-115-12

Query Match 2.0%; Score 9; DB 3; Length 502;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
|||||
DB 327 HRDLKSRNV 335

RESULT 22

US-08-436-265-12
; Sequence 12, Application US/08436265
; Patent No. 6316217
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohel; Dijke, Peter Ten;
; APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
; TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,265
; FILING DATE: 30-October-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02367
; FILING DATE: 17-NO. 6316217ember-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9224057.1
; FILING DATE: 17-NO. 6316217ember-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304677.9
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304680.3
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9311047.6
; FILING DATE: 28-May-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9313763.6
; FILING DATE: 2-July-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9136099.2
; FILING DATE: 3-August-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9321344.5
; FILING DATE: 15-October-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohel, Vineet
; REGISTRATION NUMBER: 37,003
; REFERENCE/DOCKET NUMBER: LUD 5298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-436-265-12

Query Match 2.0%; Score 9; DB 4; Length 502;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
|||||
DB 327 HRDLKSRNV 335

RESULT 23

US-09-679-187-12
; Sequence 12, Application US/09679187
; Patent No. 6331621
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohel; Dijke, Peter Ten;
; APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
; TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/679,187
; FILING DATE: 03-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,265
; FILING DATE: 30-October-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02367
; FILING DATE: 17-NO. 6331621ember-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9224057.1
; FILING DATE: 17-NO. 6331621ember-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304677.9
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9304680.3
; FILING DATE: 8-March-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9311047.6
; FILING DATE: 28-May-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9313763.6
; FILING DATE: 2-July-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9136099.2
; FILING DATE: 3-August-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9321344.5
; FILING DATE: 15-October-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohel, Vineet
; REGISTRATION NUMBER: 37,003
; REFERENCE/DOCKET NUMBER: LUD 5298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-679-187-12

Query Match 2.0%; Score 9; DB 4; Length 502;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
DB 327 HRDLKSRNV 335

RESULT 24
US-08-149-105-16
Sequence 16, Application US/08149105

PATENT INFORMATION:
PATENT NO. 5538892
APPLICANT: Donahoe, Patricia K.
APPLICANT: Gustafson, Michael
APPLICANT: He, Wei W.
APPLICANT: Wang, Xiao-Fan
TITLE OF INVENTION: TGF- TYPE I RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,105
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,673
FILING DATE: March 11, 1993
APPLICATION NUMBER: 07/853,396
FILING DATE: March 18, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/211001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 505
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-149-105-16

Query Match 2.0%; Score 9; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
DB 330 HRDLKSRNV 338

RESULT 25
US-08-317-847-16
Sequence 16, Application US/08317847

PATENT No. 5547854
GENERAL INFORMATION:

APPLICANT: Donahoe, Patricia K.
APPLICANT: Gustafson, Michael
APPLICANT: He, Wei W.
TITLE OF INVENTION: FOUR NOVEL RECEPTORS OF THE TGF-B
TITLE OF INVENTION: FAMILY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,847
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,673
FILING DATE: March 11, 1993
APPLICATION NUMBER: 07/853,396
FILING DATE: March 18, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/127002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 505
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-317-847-16

Query Match 2.0%; Score 9; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139
DB 330 HRDLKSRNV 338

RESULT 26
US-09-699-266A-13
Sequence 13, Application US/09699266A
PATENT No. 6559354
GENERAL INFORMATION:
APPLICANT: Ma, Hongchang
APPLICANT: Morakinyo, Layo O.
APPLICANT: Odell, Joan T.
APPLICANT: Orozco Jr., Emil M.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: TRANSCRIPTION AND GENE EXPRESSION REGULATORS
FILE REFERENCE: BB164 US NA
CURRENT APPLICATION NUMBER: US/09/699,266A
FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: PCT/US99/08385
PRIOR FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/083,212
PRIOR FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 13

SOFTWARE: Microsoft Office 97
SEQ ID NO: 13
LENGTH: 856
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-699-266A-13

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 856;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447
Db 17 DDDDDGEE 25

RESULT 27
US-08-446-855A-2
Sequence 2, Application US/08446855A
Patent No. 5848573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchell, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2391 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-855A-2

Query Match
Best Local Similarity 100.0%; Score 9; DB 2; Length 2391;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 445
Db 1786 DDDDDGEE 1794

RESULT 28
US-09-150-741-2
Sequence 2, Application US/09150741
Patent No. 6183996
GENERAL INFORMATION:

APPLICANT: Stewart et al.
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
Patent No. 6183996
TITLE OF INVENTION: Synthetase II
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/150,741
CURRENT FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: PL6380
EARLIER FILING DATE: 1992-12-16
EARLIER APPLICATION NUMBER: A093/00617
EARLIER FILING DATE: 1993-12-02
EARLIER APPLICATION NUMBER: 08/446,855
EARLIER FILING DATE: 1995-07-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 2
LENGTH: 2391
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2

Query Match
Best Local Similarity 100.0%; Score 9; DB 3; Length 2391;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 445
Db 1786 DDDDDGEE 1794

RESULT 29
US-08-240-712-35
Sequence 35, Application US/08240712
Patent No. 5539907
GENERAL INFORMATION:
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHERS, ANTONY JAMES
APPLICANT: STETLER, GARY L.
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
TITLE OF INVENTION: HEMOGLOBINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,712
FILING DATE: 09-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09752
FILING DATE: 13-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: ANDERSON=6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: /note= one or both of Gly
OTHER INFORMATION: residues 2 and 3 can be absent; any or all
OTHER INFORMATION: of Asp residues 5-33 can be absent; one
OTHER INFORMATION: or both of Gly residues 35 and 36 can be
OTHER INFORMATION: absent
MOLECULE TYPE: peptide
US-08-240-712-35

Query Match 1.8%; Score 8; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 3 GDDDDDD 10

RESULT 30
US-08-443-890-35
; Sequence 35, Application US/08443890
; Patent No. 5739011
; GENERAL INFORMATION:
; APPLICANT: ANDERSON, DAVID C.
; APPLICANT: MATHEWS, ANTONY JAMES
; APPLICANT: STETLER, GARY L.
; TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: US/08/443,890
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/240,712
; FILING DATE: 09-MAY-1994
; APPLICATION NUMBER: PCT/US92/09752
; FILING DATE: 13-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: ANDERSON=6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: /note= one or both of Gly
; OTHER INFORMATION: residues 2 and 3 can be absent; any or all
; OTHER INFORMATION: of Asp residues 5-33 can be absent; one
; OTHER INFORMATION: or both of Gly residues 35 and 36 can be
; OTHER INFORMATION: absent

MOLECULE TYPE: peptide
US-08-443-890-35

Query Match 1.8%; Score 8; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 3 GDDDDDD 10

RESULT 31
US-09-303-814-1
; Sequence 1, Application US/09303814
; Patent No. 6391848
; GENERAL INFORMATION:
; APPLICANT: de Lumen, Benito O.
; APPLICANT: Galvez, Alfredo F.
; TITLE OF INVENTION: Soybean Protein Nutraceuicals
; FILE REFERENCE: B99-089
; CURRENT APPLICATION NUMBER: US/09/303,814
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: soybean
US-09-303-814-1

Query Match 1.8%; Score 8; DB 4; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 34 GDDDDDD 41

RESULT 32
US-08-618-911-4
; Sequence 4, Application US/08618911
; Patent No. 5858016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; NUMBER OF SEQUENCES: 13
; TITLE OF INVENTION: SEEDS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-618-911-4

Query Match 1.8%; Score 8; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 55 GDDDDDD 62

RESULT 33
US-08-618-911-6
; Sequence 6, Application US/08618911
; Patent No. 5850016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; TITLE OF INVENTION: SEEDS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-618-911-6

Query Match 1.8%; Score 8; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 55 GDDDDDD 62

RESULT 34
US-08-938-675A-2
; Sequence 2, Application US/08938675A
; Patent No. 6107287

GENERAL INFORMATION:
; APPLICANT: de Lumen, Benito O.
; APPLICANT: Galvez, Alfredo F.
; TITLE OF INVENTION: Luminasin Peptides
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,675A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B98-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-938-675A-2

Query Match 1.8%; Score 8; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443
Db 55 GDDDDDD 62

RESULT 35
US-09-531-727-2
; Sequence 2, Application US/09531727
; Patent No. 6544956
; GENERAL INFORMATION:
; APPLICANT: de Lumen, Benito O.
; APPLICANT: Galvez, Alfredo F.
; TITLE OF INVENTION: Luminasin Peptides
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/531,727
; FILING DATE: 21-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/938,675

```

; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: OSMAN, RICHARD A
;   REGISTRATION NUMBER: 36,627
;   REFERENCE/DOCKET NUMBER: B98-003
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (650) 343-4341
;     TELEFAX: (650) 343-4342
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 158 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-531-727-2

Query Match      1.8%; Score 8; DB 4; Length 158;
Best Local Similarity 100.0%; Pred.No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      436 GDDDDDD 443
Db      55 GDDDDDD 62

RESULT 36
US-08-752-132-2
; Sequence 2, Application US/08752132
; Patent No. 5869290
; GENERAL INFORMATION:
;   APPLICANT: Freeman, Kathryn
;   APPLICANT: Nicholas, Richard
;   TITLE OF INVENTION: CAYAE1
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSER: Smithkline Beecham Corporation
;     STREET: 709 Swedeland Road
;     CITY: King of Prussia
;     STATE: PA
;     COUNTRY: U.S.A.
;     ZIP: 19406-0939
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FASTSEQ Version 1.5
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/752,132
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Gimmil, Edward R
;     REGISTRATION NUMBER: 38,891
;     REFERENCE/DOCKET NUMBER: P50580
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 610-270-4478
;       TELEFAX: 610-270-5090
;     TELEX:
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 179 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     HYPOTHETICAL: NO
;     ANTI-SENSE: NO

```

```

; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-752-132-2

Query Match      1.8%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred.No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      436 GDDDDDD 443
Db      117 GDDDDDD 124

RESULT 37
US-09-382-080-1
; Sequence 1, Application US/09382080
; Patent No. 6087333
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Goli, Surya K.
;   TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;     ADDRESSER: Incyte Pharmaceuticals, Inc.
;     STREET: 3174 Porter Drive
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FASTSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/382,080
;   FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/859,937
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Billings, Lucy J.
;     REGISTRATION NUMBER: 36,749
;     REFERENCE/DOCKET NUMBER: PF-0308 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415-855-0555
;       TELEFAX: 415-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 192 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: THPINOT03
;       CLONE: 2448450
;     US-09-382-080-1

Query Match      1.8%; Score 8; DB 3; Length 192;
Best Local Similarity 100.0%; Pred.No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      436 GDDDDDD 443
Db      118 GDDDDDD 125

RESULT 38
US-08-859-937-1
; Sequence 1, Application US/08859937
; Patent No. 6090577
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.

```

```

; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,937
; FILING DATE: Filed Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0308 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: THPINOT03
; CLONE: 2448450
; US-08-859-937-1

Query Match      1.8%; Score 8; DB 3; Length 192;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      436 GDDDDDD 443
DB      118 GDDDDDD 125

RESULT 39
; Sequence 26, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,080
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEEX: 79-0924
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-431-080-26

Query Match      1.8%; Score 8; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      437 DDDDDDDG 444
DB      142 DDDDDDDG 149

RESULT 40
; Sequence 26, Application US/08938534
; Patent No. 5916752
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,534
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEEX: 79-0924
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

US-08-938-534-26

Query Match 1.8%; Score 8; DB 2; Length 226;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 142 DDDDDDDG 149

RESULT 41

US-09-345-294-26

Sequence 26, Application US/09345294

Patent No. 6387619

GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.

Singer, Miriam S.

TITLE OF INVENTION: Telomerase Compositions and Methods

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TEXAS

COUNTRY: UNITED STATES OF AMERICA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/345,294

FILING DATE: 30-Jun-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/431,080

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 226 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Query Match 1.8%; Score 8; DB 4; Length 226;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 142 DDDDDDDG 149

RESULT 42

US-08-469-537A-80

Sequence 80, Application US/08469537A

Patent No. 5843749

GENERAL INFORMATION:

APPLICANT: Maisompierre, et al.

TITLE OF INVENTION: EHK AND ROR TYROSINE

TITLE OF INVENTION: KINASES

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill River Road

CITY: Tarrytown

STATE: NY

COUNTRY: U.S.A.

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/406,247

FILING DATE: 17-MAR-1995

APPLICATION NUMBER: USSN 08/144,992

FILING DATE: 28-OCT-1993

APPLICATION NUMBER: USSN 07/736,559

FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kempner, Ph.D., Gall M

REGISTRATION NUMBER: 32,143

REFERENCE/DOCKET NUMBER: REG 070C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-7400

TELEFAX: 914-345-7721

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

QY 191 SYGVVIME 198

Db 109 SYGVVIME 116

RESULT 43

US-08-469-537A-81

Sequence 81, Application US/08469537A

Patent No. 5843749

GENERAL INFORMATION:

APPLICANT: Maisompierre, et al.

TITLE OF INVENTION: EHK AND ROR TYROSINE

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill River Road

CITY: Tarrytown

STATE: NY

COUNTRY: U.S.A.

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 81:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-81

Query Match 1.8%; Score 8; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVLTWE 198
Db 109 SYGVLTWE 116

RESULT 44
US-09-252-991A-28073
Sequence 28073, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28073
LENGTH: 265
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28073

Query Match 1.8%; Score 8; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PLPLPLAA 336
Db 137 PLPLPLAA 144

RESULT 45
US-08-852-743-5
Sequence 5, Application US/08852743
Patent No. 5830699
GENERAL INFORMATION:
APPLICANT: Force, Thomas
APPLICANT: Kyriakis, John M.
APPLICANT: Pombo, Celia M.

APPLICANT: Bonventure, Joseph
TITLE OF INVENTION: SOK-1 AND METHODS OF USE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,743
FILING DATE: 7-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/016,774
FILING DATE: 7-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00786/327001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-743-5

Query Match 1.8%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 165 WMAPEVIO 172

RESULT 46
US-09-185-370-5
Sequence 5, Application US/09185370
Patent No. 6093560
GENERAL INFORMATION:
APPLICANT: Force, Thomas
APPLICANT: Kyriakis, John M.
APPLICANT: Pombo, Celia M.
APPLICANT: Bonventure, Joseph
TITLE OF INVENTION: SOK-1 AND METHODS OF USE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/185,370
FILING DATE:


```
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/852,743
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fraser, Janis K.
/ REGISTRATION NUMBER: 34,819
/ REFERENCE/DOCKET NUMBER: 00786/327001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 270 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-185-370-5

Query Match
Best Local Similarity 100.0%; Score 8; DB 3; Length 270;
Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 165 WMAPEVIQ 172

RESULT 47
US-09-485-077A-17
/ Sequence 17, Application US/09485077A
/ Patent No. 6458590
/ GENERAL INFORMATION:
/ APPLICANT: Mukherjee, Anil
/ APPLICANT: Kundu, Gopal
/ APPLICANT: Panda, Dibyendu
/ TITLE OF INVENTION: Methods and Compositions for Treatment of Restenosis
/ FILE REFERENCE: NIH-05047
/ CURRENT APPLICATION NUMBER: US/09/485,077A
/ CURRENT FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: PCT/US98/16569
/ PRIOR FILING DATE: 1998-07-08
/ PRIOR APPLICATION NUMBER: 60/054,967
/ PRIOR FILING DATE: 1997-07-08
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 17
/ LENGTH: 317
/ TYPE: PRT
/ ORGANISM: Rattus No. 645850vegicus
/ US-09-485-077A-17

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 317;
Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 88 DDDDDDDG 95

RESULT 48
US-09-345-473E-18
/ Sequence 18, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
```

```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18
/ LENGTH: 328
/ TYPE: PRT
/ ORGANISM: C. elegans
/ US-09-345-473E-18

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 328;
Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILS 58
DB 90 EKEAEILS 97

RESULT 49
US-09-211-930-3
/ Sequence 3, Application US/09211930
/ Patent No. 5962265
/ GENERAL INFORMATION:
/ APPLICANT: Tyrell E. No. 5962265tris
/ APPLICANT: William Craig Moore
/ APPLICANT: David Shay Silberstein
/ TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
/ FILE REFERENCE: PHM.70296
/ CURRENT APPLICATION NUMBER: US/09/211,930
/ CURRENT FILING DATE: 1998-12-15
/ EARLIER APPLICATION NUMBER: GB 9726851.0
/ EARLIER FILING DATE: 1997-12-19
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3
/ LENGTH: 416
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-211-930-3

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 416;
Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 185 WMAPEVIQ 192

RESULT 50
US-09-211-930-11
/ Sequence 11, Application US/09211930
/ Patent No. 5962265
/ GENERAL INFORMATION:
/ APPLICANT: Tyrell E. No. 5962265tris
/ APPLICANT: William Craig Moore
/ APPLICANT: David Shay Silberstein
/ TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
/ FILE REFERENCE: PHM.70296
/ CURRENT APPLICATION NUMBER: US/09/211,930
/ CURRENT FILING DATE: 1998-12-15
/ EARLIER APPLICATION NUMBER: GB 9726851.0
/ EARLIER FILING DATE: 1997-12-19
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 416
/ TYPE: PRT
/ ORGANISM: Mus musculus
/ US-09-211-930-11

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 416;
Pred. No. 32;
```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 51
US-09-340-993-3
; Sequence 3, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-993-3

Query Match
Best Local Similarity 100.0%; Score 8; DB 3; Length 416;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 52
US-09-340-993-11
; Sequence 11, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-993-11

Query Match
Best Local Similarity 100.0%; Score 8; DB 3; Length 416;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 53
US-09-468-442-3
; Sequence 3, Application US/09468442
; Patent No. 6300098

; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-468-442-3

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 416;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 54
US-09-468-442-11
; Sequence 11, Application US/09468442
; Patent No. 6300098
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-468-442-11

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 416;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 55
US-09-685-462-4
; Sequence 4, Application US/09685462
; Patent No. 6524833
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard I.
; APPLICANT: Agarwal, Sachana

APPLICANT: Best, Jennifer
APPLICANT: Vail, Brenda
TITLE OF INVENTION: Two Sterile-20 Kinase-Like Proteins and
FILE REFERENCE: 1242.1016-004
CURRENT APPLICATION NUMBER: US/09/685,462
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/591,083
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: PCT/US98/26116
PRIOR FILING DATE: 1998-12-09
PRIOR APPLICATION NUMBER: US 60/069,078
PRIOR FILING DATE: 1997-12-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 416
TYPE: PRT
ORGANISM: Mus musculus
US-09-685-462-4

Query Match 1.8%; Score 8; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 56
US-09-345-473E-14
Sequence 14, Application US/09345473E
Patent No. 6558903
GENERAL INFORMATION:
APPLICANT: Hodges, Martin
TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
FILE REFERENCE: 35800/183781
CURRENT APPLICATION NUMBER: US/09/345,473E
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 416
TYPE: PRT
ORGANISM: Homo sapiens
US-09-345-473E-14

Query Match 1.8%; Score 8; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192

RESULT 57
US-08-712-709-8
Sequence 8, Application US/08712709
Patent No. 5863780
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,709
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-08-712-709-8

Query Match 1.8%; Score 8; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 190 WMAPEVIQ 197

RESULT 58
US-09-111-444-8
Sequence 8, Application US/09111444
Patent No. 6045792
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,444
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-09-111-444-8

Query Match 1.8%; Score 8; DB 3; Length 487;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
|||||
DB 190 WMAPEVIQ 197

RESULT 59
US-09-541-228-8
Sequence 8, Application US/09541228
Patent No. 6232077
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-09-541-228-8

Query Match 1.8%; Score 8; DB 3; Length 487;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 WMAPEVIQ 179
|||||

DB 190 WMAPEVIQ 197

RESULT 60
US-09-685-462-8
Sequence 8, Application US/09685462
Patent No. 6524833
GENERAL INFORMATION:
APPLICANT: Zon, Leonard I.
APPLICANT: Agarwal, Sadhana
APPLICANT: Best, Jennifer
APPLICANT: Vail, Brenda
TITLE OF INVENTION: Two Sterile-20 Kinase-Like Proteins and
TITLE OF INVENTION: Methods of Use Thereof
FILE REFERENCE: 1242.1016-004
CURRENT APPLICATION NUMBER: US/09/685,462
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/591,083
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: PCT/US98/26116
PRIOR FILING DATE: 1998-12-09
PRIOR APPLICATION NUMBER: US 60/069,078
PRIOR FILING DATE: 1997-12-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 487
TYPE: PRT
ORGANISM: Homo sapiens
US-09-685-462-8

Query Match 1.8%; Score 8; DB 4; Length 487;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
|||||
DB 190 WMAPEVIQ 197

RESULT 61
US-09-252-991A-26096
Sequence 26096, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26096
LENGTH: 544
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26096

Query Match 1.8%; Score 8; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 EOSNTPPL 331
|||||
DB 370 EOSNTPPL 377

RESULT 62
US-09-529-279-4

```
/ Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match      1.8%; Score 8; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 VLKICDFG 153
Db      170 VLKICDFG 177

RESULT 63
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match      1.8%; Score 8; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 VLKICDFG 153
Db      170 VLKICDFG 177

RESULT 64
US-09-529-279-15
; Sequence 15, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
```

```
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-15

Query Match      1.8%; Score 8; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 VLKICDFG 153
Db      170 VLKICDFG 177

RESULT 65
US-10-158-895-15
; Sequence 15, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-15

Query Match      1.8%; Score 8; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 VLKICDFG 153
Db      170 VLKICDFG 177

RESULT 66
US-08-469-537A-107
; Sequence 107, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisondierre, et al.
; TITLE OF INVENTION: BHK AND ROR TYROSINE
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
```

CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Human ROR2
LOCATION: 1...943
OTHER INFORMATION:
US-08-469-537A-107

Query Match 1.8%; Score 8; DB 2; Length 943;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVWLME 198
Db 677 SYGVWLME 684

RESULT 67
US-08-431-080-28
Sequence 28, Application US/0841080
Patent No. 5698686
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/431,080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-28

Query Match 1.8%; Score 8; DB 1; Length 1085;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 142 DDDDDDDG 149

RESULT 68
US-08-938-534-28
Sequence 28, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,534
FILING DATE: 26-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-534-28

Query Match 1.8%; Score 8; DB 2; Length 1085;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 142 DDDDDDDG 149

RESULT 69
US-09-345-294-28
Sequence 28, Application US/09345294
Patent No. 6387619
GENERAL INFORMATION:
APPLICANT: Gortschling, Daniel E.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/345,294
FILING DATE: 30-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-345-294-28

Query Match 1.8%; Score 8; DB 4; Length 1085;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 142 DDDDDDDG 149

RESULT 70
US-08-482-228-215
Sequence 215, Application US/08482228
Patent No. 5968753
GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet
APPLICANT: Kobori, Joan A.
APPLICANT: Al-Abdaly, Fahad A.
APPLICANT: Guillermo, Roy
APPLICANT: Helgerson, Sam L.
APPLICANT: Deans, Robert J.
TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL
SELECTION MEDIATED BY PEPTIDE RELEASE
NUMBER OF SEQUENCES: 215
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janice Guthrie, Ph.D.
STREET: P.O. Box 15210
CITY: Irvine
STATE: California
COUNTRY: USA
ZIP: 92713-5210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,228
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Guthrie, Janice
REGISTRATION NUMBER: 35,170
REFERENCE/DOCKET NUMBER: IT-4630CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 440-5353
TELEFAX: (714) 553-1952
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-228-215

Query Match 1.5%; Score 7; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 1 DDDDDDD 7

RESULT 71
US-08-482-528-215
Sequence 215, Application US/08482528
Patent No. 6017719
GENERAL INFORMATION:
APPLICANT: Tseng-Law, Janet
APPLICANT: Kobori, Joan A.
APPLICANT: Al-Abdaly, Fahad A.
APPLICANT: Guillermo, Roy
APPLICANT: Helgerson, Sam L.
APPLICANT: Deans, Robert J.
TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL
SELECTION MEDIATED BY PEPTIDE RELEASE
NUMBER OF SEQUENCES: 215
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janice Guthrie, Ph.D.
STREET: P.O. Box 15210
CITY: Irvine
STATE: California
COUNTRY: USA
ZIP: 92713-5210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,528
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Guthrie, Janice
REGISTRATION NUMBER: 35,170
REFERENCE/DOCKET NUMBER: IT-4630CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 440-5353
TELEFAX: (714) 553-1952
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-528-215

Query Match 1.5%; Score 7; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
DB 1 DDDDDD 7

RESULT 72
US-08-199-776-23
Sequence 23, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120e1 integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
City: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: synthetic peptide

US-08-199-776-23

Query Match 1.5%; Score 7; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
DB 3 DDDDDD 9

RESULT 73
US-08-663-731-23
Sequence 23, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423e1 integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
City: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: synthetic peptide
US-08-663-731-23

Query Match 1.5%; Score 7; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
DB 3 DDDDDD 9

RESULT 74
US-08-879-338-23
Sequence 23, Application US/08879338A
Patent No. 6063906
GENERAL INFORMATION:


```
/ APPLICANT: Brenner, Michael B.
/ APPLICANT: Parker, Christina M.
/ TITLE OF INVENTION: Antibodies to No. 6063906e1 Integrin Alpha
/ TITLE OF INVENTION: Subunit
/ FILE REFERENCE: B0801/7080/ERP
/ CURRENT APPLICATION NUMBER: US/08/879,338A
/ CURRENT FILING DATE: 1997-06-20
/ EARLIER APPLICATION NUMBER: US 08/663,731
/ EARLIER FILING DATE: 1996-06-14
/ EARLIER APPLICATION NUMBER: US 08/199,776
/ EARLIER FILING DATE: 1994-02-18
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 23
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic Peptide
US-08-879-338-23
```

```
Query Match 1.5%; Score 7; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 437 DDDDDDD 443

Db 3 DDDDDDD 9

```
RESULT 75
US-09-293-238B-23
/ Sequence 23, Application US/09293238B
/ Patent No. 6455042
/ GENERAL INFORMATION:
/ APPLICANT: Brenner, Michael B.
/ APPLICANT: Parker, Christina M.
/ TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
/ TITLE OF INVENTION: or Crohn's Disease by Administering an Antibody to Alpha E
/ TITLE OF INVENTION: Beta 7 Integrin
/ FILE REFERENCE: L0560/7005/ERP
/ CURRENT APPLICATION NUMBER: US/09/293,238B
/ CURRENT FILING DATE: 1999-04-16
/ PRIOR APPLICATION NUMBER: US 08/879,338
/ PRIOR FILING DATE: 1997-06-20
/ PRIOR APPLICATION NUMBER: US 08/663,731
/ PRIOR FILING DATE: 1996-06-14
/ PRIOR APPLICATION NUMBER: US 08/199,776
/ PRIOR FILING DATE: 1994-02-18
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 23
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic Peptide
US-09-293-238B-23
```

```
Query Match 1.5%; Score 7; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 437 DDDDDDD 443

Db 3 DDDDDDD 9

Search completed: December 5, 2003, 09:32:47
Job time : 24 secs

THIS PAGE BLANK (USPTO)